DOCUMENT RESUME

ED 216 516 FL 012 869

AUTHOR Piper, Richard M; And Others
TITLE Improving Estimates of Achievement in Students with Limited English Proficiency.
INSTITUTION Applied Social Research, Inc., Pacific Palisades, CA.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 81
GRANT NIE-G-80-0081
NOTE 129p.

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Bilingual Students; Elementary Secondary Education; *Limited English Speaking; Spanish Speaking; *Standardized Tests; Testing Programs; *Test Wiseness

ABSTRACT

The majority of Latin American immigrant students in ESEA Title VII Spanish/English programs that have been evaluated seem to suffer from inexperience in taking tests like the Comprehensive Tests of Basic Skills (CTBS). This study was designed to assess test familiarity, language of the test, and literacy, and to examine whether certain alterations in the testing procedure might allow for better estimation of students' ability and program benefits. Limited English proficiency students in grades four, six, and eight participated in the study. Instruments used were the English and Spanish versions of the CTBS. Testwiseness training was developed and administered to half the groups. An oral administration procedure for students with limited native language literacy skills as well as limited English proficiency was also developed. A sociolinguistic analysis and a psychometric analysis of the tests were developed and testwiseness materials were designed. The history of the development of the procedures, as well as the tryouts and revisions, a description of the teacher training, and a detailed analysis of the testing data comprise the greater part of the report. Attached to the study is the English version of a teacher's manual for training in testwiseness, a program adapted for use with the CTBS. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216516

IMPROVING ESTIMATES OF ACHIEVEMENT IN STUDENTS WITH LIMITED ENGLISH PROFICIENCY

A Study Performed Under a Grant Award
from the National Institute of Education
Grant Number NIE-G-80-0081

Principal Investigator

Richard M. Piper

Investigators

William J. Doherty
Nancy A. Russo

Applied Social Research, Inc.
15219 Sunset Boulevard, Suite 201
Pacific Palisades, California 90272
(213) 454-8464

FL012869

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TABLE OF CONTENTS

Page

1. The Problem. 1

2. Design of the Study. 2

3. Procedures 4

4. Results. 24

   A. The linguistic analysis. 26

   B. An analysis of effects 59

5. Discussion and conclusions 72

6. Appendix: "Teacher's Manual for Training in Testwiseness 81

ACKNOWLEDGEMENTS

The authors wish to express their appreciation for invaluable assistance in the conduct of this research. Thanks go to:

(1) Belia Cruz, Arden Daniels, Sabiniano Legarreta, Tom Martin, Alice Petrossian, Dr. Martin Pilgreen, Dr. Marie Plakos, Elsa Santillan, Mary Skandera, and Lilia Stapleton for permission to conduct the study in their respective districts and schools.

(2) Linda Calderon, Mary Calderwood, Graciela Carmona, Hector Carrio, Suanna Ponce Gilman, Barbara Harper, and María Otero for coordinating the work in their respective districts, conducting testwiseness training and testing.

(3) Dr. Lois Weinberg for writing the English version of the testwiseness training materials.

(4) Dr. Joe Lucero for designing and assisting in the socio-linguistic analysis of the CTBS.

(5) Tony Brito, Dr. Victor Rodríguez, and Linda Smith for recording the CTBS items.

THE PROBLEM

Several years ago, the Principal Investigator (P.I.) served as a teacher in a school in Nicaragua. The school, as required by law, was organized according to the French education model. Thus, instruction was highly oriented toward the development of academic skills. High performance was required. Those who could not meet the standards either dropped out or flunked out. The few who made it into and through the bachillerato program were the intellectual cream of the country.

Since many of these bachilleres would be going to college in the United States, the P.I. became interested in determining their academic status vis a vis high school graduates in the United States. Permission was obtained to give them the Spanish version of the Differential Aptitude Test (DAT) battery. The expectation was that, based on their performance in school, they would obtain scores in the upper ranges of the U.S. norm distribution. This expectation was not fulfilled. In fact the students did quite poorly. This unexpected result led the P.I. to conduct a discussion with the students. One thing that seems to have bothered them was the format of the test. It was different than anything they had experienced in their educational careers. As a consequence, they had never developed the test taking skills relevant specifically to taking objective norm-referenced tests like the DAT.

Since returning to the United States, the P.I. has served as the evaluator of a number of ESEA Title VII Spanish/English bilingual programs. The majority of Latin American immigrant students in the programs evaluated seem to suffer from the same inexperience in taking tests like the Comprehensive Tests of Basic Skills (CTBS). If this is true, then in all probability, the test scores they obtain underestimate their true levels of performance.

There may be other problems in using tests like the CTBS to measure the educational achievement of language minority students. One of these is the language of the test. Prior to the introduction of the CTBS Espanol, many of these students were examined using the CTBS, Form S (CTBS/S). The absence of a match between the language of the test and the language of the examinee as well as the absence of a match between the oral/aural language tasks addressed in English as a Second Language (ESL) programs and the language tasks addressed in the CTBS/S motivated recommendations of caution to client school districts regarding the use of the test with language minority students and, when used, the interpretation of scores. It was felt that such scores systematically underestimated both student conceptual skills and program benefits.

Given these feelings of discomfort with the CTBS/S, the introduction of the CTBS Espanol was welcomed. This version, which was largely a translation of the CTBS/S (some unique Spanish language material was introduced for linguistic and cultural reasons), began to be used by some districts with some Spanish dominant students. The results of this shift were predictable. These students, depending on their levels of primary language literacy, regularly obtained higher scores on the reading, language and mathematics concepts and applications subtests than they had previously obtained on the parallel English language version. The CTBS Espanol offered a means for estimating student conceptual skills as well as the benefits of instruction in Spanish reading and language arts.

Questions still remained. Many practitioners questioned the validity of a Spanish-language test that was merely a translation of an English-language test. To what degree would such a test reflect non-Spanish modes of expression? To what degree would it be culturally biased? Further, what inferences could be made based on item scores? If a student failed an item, could this be attributed merely to an absence of literacy or was there, beyond a possible absence of literacy, a lack of knowledge of the concepts represented by the words?

Out of these observations and questions, the problem addressed by this study was formulated. The major assumption underlying the study is that the factorial composition of tests like the CTBS/S is different for students with limited English proficiency (LEP) than for students with full English proficiency (FEP). This factorial complexity introduces systematic error into estimates of LEP students' ability because the score is not a pure measure of the ability. The sources of this factorial complexity may be many. This study focuses on three: unfamiliarity with the test format, literacy, and the language of the test.

The study, then, was designed to directly assess test familiarity, language of the test, and literacy and to examine whether certain alterations in the testing procedure might allow for better estimation of students' ability and program benefits.

DESIGN OF THE STUDY

A three-factor factorial design was used. The factors were (1) language of the test (Spanish, English), (2) administration conditions (standard, oral) and (3) test familiarity (trained, untrained). By crossing these factors, eight treatment groups were obtained as follows:

1. Spanish version of the test, standard administration, testwiseness training.

2. Spanish version of the test, standard administration, no testwiseness training.
3. Spanish version of the test, oral administration, testwiseness training.
4. Spanish version of the test, oral administration, no testwiseness training.
5. English version of the test, standard administration, testwiseness training.
6. English version of the test, standard administration, no testwiseness training.
7. English version of the test, oral administration, testwiseness training.
8. English version of the test, oral administration, no testwiseness training.

Students who participated in the study came from grades four, six, and eight. All had Spanish as their primary language and were limited English proficient. In order to control for factors that might bias the result, they were assigned to conditions on a modified matching basis in which controls were instituted for country of birth, amount of prior education, socio-economic status, length of residence in the U.S., grade level, and entering levels of English and Spanish language proficiency.

Instruments used were the CTBS, Form S and the CTBS Espanol, Level 1 (with fourth grade participants) and Level 3 (with sixth and eighth grade participants). The latter was a translation of the former completed by staff in the Norwalk-La Mirada (CA) Unified School District. In some instances, material that was considered inappropriate for translation was replaced by more culturally and/or linguistically appropriate material.

The testwiseness training that was administered to half the groups was developed as a part of this study. Its characteristics are described later in this report.

The oral administration procedure was also developed as a part of this study. Its characteristics are also described later in this report.

PROCEDURES

Linguistic Analysis of the CTBS

It is frequently claimed that norm-referenced tests used for measuring the aptitude and achievement of students in the United States are culturally biased. The claim is accepted as true by a large number of educators and, in response to the claim, a variety of attempts have been made to create culture-free or culture-fair tests. Indeed, any manufacturer who produces and markets a test for broad use must demonstrate that some effort to rid the instrument of culture bias has been made.

Such is the case with the manufacturer of the CTBS. They define bias as the systematic production of unfair scores for a particular group. In the technical manual, the manufacturer carefully describes the item tryout and evaluation procedure with samples of Black, Spanish-speaking, "other," and "standard" students. In the judgment of the writers, reasonable care was taken to eliminate bias as defined.

Nevertheless, for all the care that may be taken, it is impossible to remove all bias from a test. This follows from the fact that item writer(s) write from within the framework of their own culture(s). They can do no other. There is no such thing as a culture-free person or a culture-free position. This being true, any test will be unbiased for some populations and biased for others. We cannot remove bias from a given test for all prospective populations of examinees. We can only replace one bias with some other bias. Given, therefore, that bias will inevitability be present, our task is to try to reduce its effects for those sub-populations that are negatively affected. Our first step in this direction was a comparative socio-linguistic analysis of the English and Spanish versions of the CTBS.

Design of the Linguistic Analysis

Two analyses were designed, a socio-linguistic analysis and a psychometric analysis. The socio-linguistic analysis was designed to facilitate a direct comparison of the English and Spanish versions as well as to inform the interpretation of our item analysis. Accordingly, parallel English and Spanish versions of items from the Reading Vocabulary, Reading Comprehension, Mathematics Computation, and Mathematics Concepts and Applications subtests were pasted on the face of 5X8 index cards. On the back of each of these cards was pasted the analytic outline presented in Exhibit 1.

EXHIBIT 1

OUTLINE FOR CONDUCTING THE
SOCIO-LINGUISTIC ITEM ANALYSIS

Phonological interferences:

Spanish _____

English _____

Lexicon: interferences, translations, high and low frequency items:

Spanish _____

English _____

Grammatical structures:

Spanish _____

English _____

Visual cues/stimuli used with test items:

Spanish _____

English _____

Socio-linguistic variables:

Hispanic/Spanish dominant other than Mexican American _____

Hispanic/Mexican American _____

English dominant

Anglo _____

Hispanic other than Mexican American _____

Mexican American _____

Additional Comments:

_____

_____

The psychometric analysis was designed to focus on item difficulty, item discrimination, and item response frequencies.

Conduct of the Linguistic Analysis

The item cards described above were given to a Spanish/English bilingual linguist. He conducted an analysis of each Spanish/English item pair for Levels B, C, 1, 2, and 3 of the CTBS. His comments were used as a basis for interpreting the results obtained in the item analysis.

The psychometric analysis focused primarily on the individual items. Item difficulty was computed as percent of correct response. This was the form of the index used by the manufacturer and permitted a direct comparison of the difficulty indices obtained from the manufacturer's standardization sample with those obtained from the language minority sample of this study.

Item discrimination was computed in terms of point biserial correlations between each item and the rest of the items in the subtest with that item removed.

Overall test reliability was computed in terms of Cronbach's alpha.

Development of the Testwiseness Materials

Rationale

Among the factors that are assumed to contribute to systematic negative error in the test scores of language minority students is a lack of testwiseness, that is, the lack of specific skills that could assist a student to receive a higher score than he might otherwise have received. The lack of these skills reflects a lack of experience with norm-referenced tests in the home country.

In order to provide the language minority student with those skills that he/she would have developed had he/she been born and raised in this country, a testwiseness training system was developed and used in this study.

Design and Development

The construct of testwiseness has a relatively short history in educational research. Both educational measurement specialists and test constructors have testified to its

existence. It has been described both as having a possible influence on reliability (Thorndike, 1951) and as a component of response variance in objective test questions (Ebel and Damrin, 1960). The literature reflected considerable confusion over the precise components and functioning of testwiseness. No empirical research to clarify the situation was reported prior to 1964.

The classic theoretical work was performed by Millman, Bishop, and Ebel (1965). They presented a comprehensive taxonomy of testwiseness that was intended to serve as a framework for future empirical study. They defined testwiseness as "a subject's capacity to utilize the characteristics and formats of the test and/or test-taking situation to receive a high score." Subsequent to this work, a large amount of empirical research appeared in the professional literature. Within the last decade, testwiseness has been investigated in terms of its components (Nilsson and Wedman, 1974), correlates (Diamond and Evans, 1972), and application (Ford, 1973). A comprehensive review of investigations of testwiseness has recently been conducted by Sarnacki (1979). As a result of these and other analyses, testwiseness is now widely recognized as a source of additional error variance in test scores and as a possible depressor of test validity.

Our first step in the development of materials to teach testwiseness was to conduct an analysis of the specific testwiseness skills required by LEP students in taking a standardized achievement test, in this case, the CTBS. This analysis involved three tasks: a review and identification of relevant testwiseness skills contained in the taxonomy of testwiseness skills prepared by Millman et al. (1965), a task analysis of the skills required in the sub-tests of the CTBS, and discussion with experts in educational measurement and bilingual education about the problems faced by students generally in taking standardized achievement tests, and those faced by LEP students, in particular.

Based on the above analysis, a conceptual framework for testwiseness training of LEP students was developed (see Table 1). Basically, three categories of skills were identified as most important given the nature and purpose of the study: familiarity with standardized tests, following directions, and strategies for test-taking.

Our second step was to examine previously developed materials designed to teach testwiseness skills to determine whether they would be appropriate for use in this study. These materials were located through reports of studies published in professional journals, doctoral dissertations, and contacts with school districts across the nation reputed to be training students in testwiseness skills.

Our objective in reviewing these materials was to determine whether they addressed the skills contained in our

TABLE 1

CONCEPTUAL FRAMEWORK FOR TESTWISENESS TRAINING
OF LEP STUDENTS

Familiarity with Standardized Tests

1. Features of standardized tests
2. Feelings when taking tests
3. Marking the answer sheet or test booklet quickly and properly

Following Directions

1. Reading every word
2. Understanding types of questions and answers required for each sub-test
   - Reading Vocabulary
   - Reading Comprehension
   - Mathematics Computation
   - Mathematics Concepts & Applications

   } CTBS Espanol

   - Spelling
   - Language Mechanics
   - Language Expression
   - Reference Skills

   } CTBS English

Test-Taking Strategies

1. Strategies for answering each type of question in each subtest
2. Guessing
3. Marking numbers of questions for reconsideration
4. Using scratch paper
5. Checking answers

conceptual framework, whether they were appropriate for LEP students at the grade levels identified for the study (4th, 6th, 8th grade), and whether they were usable within the amount of instructional time available for the study (5 one-hour sessions). A summary of the key characteristics of all materials located is presented in Table 2. As can be seen from the table, the majority of materials addressed more sophisticated and complex skills and were directed to older English speaking students. Based on this review, it was decided that no existing materials were appropriate for use in this study.

Thus, we began the process of developing testwiseness materials for the study. A preliminary set of materials was developed containing instructional strategies and exercises keyed to the CTBS that were appropriate for LEP students at grades four, six, and eight. The materials presented instruction on the skills identified in Table 1 in a logical sequence with activities that were designed to motivate and inform students.

The materials were divided into five sessions. The organizational format for each session was as follows: set of lessons, objectives, general ideas, activities, practice exercise, correct answers, and final assessment. The preliminary materials were developed first in English and then translated into Spanish.

Tryout

Upon completion of development of the materials, the student materials were translated into Spanish and were produced in small quantities. A tryout was then scheduled with fourth, sixth, and eighth grade monolingual Spanish speakers. The students came from three different schools, one grade level per school. Care was taken to make sure that they would exhibit the same characteristics as those exhibited by the experimental sample. Class sizes ranged from eight to 12 students. Four one-hour training sessions were scheduled at each site on four successive days. The classes were taught by a Spanish/English bilingual instructor and were observed by the writer of the materials, a person who has good aural Spanish comprehension and is Spanish literate. The same materials and procedures were used with all three groups. The only changes were in the demeanor of the instructor who tried in each case to talk with the students at their own level. Following each one-hour session, the instructor and observer critiqued the session. The focus of these critique sessions was to generate a set of specifications for revisions of the materials.

TABLE 2

PREVIOUS STUDIES USING TESTWISENESS INSTRUCTION

Study	Skill(s)	Grade Level	Format	Instructional Time
Moore et al, (1966)	Guessing Pacing of responses	8th	Programmed Text	70-90 minutes
Langer, et al, (1973)	Using cues	college	1) script (prose) 2) programmed text	15 minutes 30 minutes
Wahlstrom & Boersma (1968)	All skills in Millman taxonomy	9th	Outline of TW principles Elaboration of selected principles Teacherwiseness with individual study	100 minutes
Oakland (1972)	Familiarise students with format (items with options, working in rows and columns) Teach common concepts (right, left, same, different) Pay attention to all options Practice marking Motivation-encourage to ask questions Use a marker Increase attention span	preschool IB (error avoidance)	30 pages of practice materials	2X a week for 6 weeks
	Cue using			

TABLE 2 (Cont.)

Study	Skill(s)		Grade Level	Format	Instructional Time
Slakter et al, (1970)	stem options		Seniors	Programmed tests	40 minutes each
	absurd options				
	similar options				
	specific determiners				
	guessing				
Callenbach (1973)	following directions	IB (error avoidance)	2nd	Instruction & Practice	8 - 30 minute lessons
	response marking				
	using time				
	guessing				
Woodley (1975)	reducing test related anxiety		Adults (occupation & licensing exams)	Workbook	4 hours
	time using				
	error avoidance				
	guessing				
	Deductive reasoning				
	Interest consideration				
	Cue using				

TABLE 2 (Cont.)

Study	Skill(s)	Grade Level	Format	Instructional Time
Ford & Weener (1980)	Test mechanics deductive reasoning	College	Lecture	3 sessions
	Cue using		Supervised study (3 booklets)	
Johngma & Warshauer (1975)	Motivation	5th/inner city & suburban	Teacher-led group discussion with written activities	1 hour
	Following directions, understanding what is read			1 hour, 15 minutes (1-2 sessions)
	Guessing			
	Using answer sheets			
	Using time			
	Test-taking conduct			
Ferrell (1977)	Deductive reasoning	high school/ college	Form 2 - a test as starting point for group discussion	1 class period (1 hour)
	Cue using			

Revision

The only revisions that were needed involved a resequencing of some of the content and activities. The revisions were completed. The testwiseness materials presented with this report are the revised version.

Development of the Oral Administration Procedures

Rationale

Most young people in Latin America do not have a chance at formal education. Nevertheless, since life in the home and community can be presumed to have educational value, we must assume a great store of practical knowledge. Unfortunately, due to a lack of literacy, these students can only reveal this store through the use of oral language.

If the purpose of testing is to find out what a student knows and if what a student knows cannot be determined using a written test, then it only makes sense to adjust the testing mode in order to accomplish the stated purpose. The objection may be raised that changing the mode of testing makes the results obtained non-comparable. This is certainly true in some, perhaps all, cases. An oral version of a reading comprehension test, for example, cannot be considered a test of reading comprehension. However, it may still be possible to learn something of value about a student's language and conceptual skills by presenting such a test in this way. And when it comes to certain other subjects (mathematics computation for example), what one learns with a test delivered orally may be very similar to what one learns from a printed test.

These considerations lead one to speculate about what might be learned with regard to the state of a student's knowledge from an oral administration of the CTBS. Consider that the form of the reading vocabulary items is as follows:

large house

(1) big
(2) long
(3) new
(4) ready

In this item form, the student is expected to choose the word that means the same or nearly the same as the underlined word. In the printed version of the test, the student must be able

to decode the words, assess their meanings, and make an appropriate selection. In the oral version of the test, only the stimulus situation changes. The student must still determine meanings and make an appropriate selection. From such a task, one can learn whether the meanings of the test words are known. The only thing one cannot determine is whether, or the degree to which, the written form of the words is known.

A similar argument can be made with respect to the reading comprehension subtest. Consider that the items present a short reading selection followed by four-choice multiple-choice items. These items test for such things as determination of word meanings from context, attention to detail, sequence of events, main idea, etc. In the print version of the test, the student must first decode the reading selection and the items, then make the appropriate response. In the oral version of the test, only the decoding part of the task is eliminated. The student must still deal with all the semantic requirements of the items. A similar argument can be made for the mathematics concepts and applications subtests.

Full recognition is given in this study to the fact that the oral approach almost certainly changes the measurement properties of the test. An assessment of these changes is included in the plan for data analysis.

The rationale, then, for designing and developing an oral administration procedure rests on the assumption that students with limited literacy skills (whether English or primary language) still have considerable knowledge which, because of the requirements of the print test, must remain unknown to us if we limit ourselves to the use of print. Only as we adjust our testing procedures to match the student's language capability can we unlock the door to his/her store of knowledge.

Design and Development

In developing the oral administration procedures, the intent was to leave the form of the test unchanged wherever possible. Under standard administration procedures, the examiner reads the directions. By slightly modifying the directions as given in the Examiner's Manual, it was felt that the task of reading the directions could best be left with the examiner. The only change would be to ignore the instruction "Do not read the item aloud" wherever it occurred. This seemed the most appropriate since the instruction always appeared with sample items and it was thought preferable to present the sample items in a fashion identical to the presentation of the items, that is orally. For this reason, the instructions were not recorded on tape.

The items for the reading vocabulary, the reading comprehension, and the mathematics concepts and applications subtests were recorded on tape. The mathematics computation subtest was not recorded since it does not involve written language. This study did not deal with the spelling, language mechanics, language expression, reference skills, science, and social studies subtests since these subtests are not included in the CTBS Espanol.

In preparing the recordings of the reading vocabulary, reading comprehension, and mathematics concepts and applications subtests, it was felt that pacing and understandability would be the most critical factors. With regard to pacing, it was decided to record each subtest such that it would fit within the time limits specified for the print version. Thus, the run time for the reading vocabulary subtest had to be less than 15 minutes for Level 1 and 12 minutes for Level 3, the reading comprehension subtest less than 35 for both Levels 1 and 3, the mathematics concepts and applications subtest less than 35 minutes for both Levels 1 and 3.

The actual run time would of course be affected by the pace of reading, the number of item repetitions, and the amount of silence between items. It was decided to read the items at a pace slightly slower than normal conversation. The other decisions varied among the test as follows:

1. On the reading vocabulary subtest, item stems were read once and response choices twice. Three seconds of silence separate the reading of each item. The total run time on the Spanish version, Level 1 is 12:50, leaving 2:10 for those students with reading skill to go back over omitted or other items. The total run time on the English version, Level 1, is 12:08, leaving 2:52. The total run time on the Spanish version, Level 3, is 11:44, leaving 0:16. The total run time on the English version, Level 3 is 11:30, leaving 0:30.
2. On the reading comprehension subtest, the "reading" selections are read aloud once. Item stems are read once and response choices twice. Three seconds of silence separate the reading of each item. The total run time on the Spanish version, Level 1 is 23:10, leaving 11:50 for those students with some reading skills to go back over omitted or other items. The total run time on the English version, Level 1 is 22:45, leaving

12:15. The total run time on the Spanish version, Level 3 is 26:00, leaving 9:00. The total run time on the English version, Level 3 is 24:11, leaving 9:49.

3. On the mathematics concepts and applications subtests, each item is read once. Response options are read only where they require the reading of words. Twenty-three seconds of silence separate the reading of each item. The total run time on the Spanish version, Level 1 is 24:30, leaving 10:30 to go back over omitted or other items. The total run time on the English version, Level 1 is 24:00. The total run time on the Spanish version, Level 3 is 24:15, leaving 10:45. The total run time on the English version, Level 3 is 24:00, leaving 11:00.

On all the recordings, the beginning of each subtest is identified with a tag giving the subtest number and title. Additionally, each item is tagged by its number in the test book, thus assisting the student to keep track of the item he/she is currently working on.

With respect to the voices used on the recordings, the Spanish version was recorded by a male college professor of Spanish whose primary language is Spanish. The English version was recorded by a female professional narrator whose primary language is English.

The technical direction and editing was performed by a professional musician with years of studio experience both as a performer and as a director/producer.

The recording was done in a professional recording studio on Ampex equipment.

Tryout and Revision

An informal tryout was conducted in order to determine whether or not the recordings needed to be revised. Of special concern was the effect of the pacing, especially the length of the silences between items. None of the evidence indicated that changes were needed though it became clear, both in the tryout setting and in the experimental setting that the time allowed is too long for some and too short for others. The fixed pacing is at once the oral administration's greatest advantage and its greatest disadvantage. The advantage lies in the pressure it places on the examinee to respond

to every item. The disadvantage lies in the degree to which it slows some students down causing them to wait and become bored, and speeds others beyond their normal pace, perhaps to their disadvantage. During tryout and testing, no examinee wanted to use the available time at the end of each subtest to answer omitted items. As a consequence, the total testing time was reduced beneath the time limits recommended by the test manufacturer.

Selection of the Sample

Problems Encountered

The proposal on which this study is based was first submitted more than three years ago. It was rejected the first year but along with the rejection came some suggestions for modification and an invitation to submit the modified proposal. The modifications were made and the invitation accepted. The time lapse between the original submission and funding of the re-submission was approximately two years. Another six months passed as the start-up work was completed. It was now time to begin preliminary work with the cooperating school districts. It was at this time that two of the three original districts decided that they could not participate. One of them had decided just that year to discontinue use of the CTBS, a key instrument in this study, and to begin using the California Achievement Test (CAT). We now faced the problem of seeking at a late hour additional districts that would cooperate with us. Such districts were found, but only at the expense of a reduction in the sample size originally planned. Later, because of internal communication problems, one of the junior high principals whose school was scheduled to supply 30 eighth grade subjects, declined to participate.

Selection of Participating Districts

As noted above, one of the three original districts continued in the project. This district (we shall designate it as District A) lies on the eastern side of a large metropolitan area adjacent to an area known as a "port of entry" for Mexican immigrants. This district has a large population of Spanish-surnamed students, a large and growing population of Asian students, and a large population of "Anglo" students.

Of the new districts, District B lies on the southwestern side of the metropolitan area and has a predominantly Black population that is being increasingly displaced on its west side by immigrant Latin American students, primarily Mexican and Central American.

District C is on the northern fringes of the large metropolitan area. It is totally different in character from Districts A and B in that it is predominantly middle to upper-middle class "Anglo" with a strong sense of pride in its traditional support of education and the arts. There is a growing immigrant Hispanic population living on the outer fringes of the city.

Districts D and E are in adjacent suburban cities lying in what may be called an "ethnic corridor" to the southeast of the large metropolitan area. District D has a large population of Mexican Americans and "Anglos." District E is ethnically diverse. District D participated fully in the study while District E agreed only to supply non-experimental control subjects.

Selection and Matching of the Students

The student participants were selected in a way designed to facilitate the formation of groups that would be as nearly alike as possible. Since we had a limited set of available students, random selection from a population was not possible. We therefore decided to approximate a matching procedure. Approximation was a necessity since strict matching is wasteful of subjects and we were in no position to give up more than a very few of the available sample.

The actual procedure required that we define a set of variables on which the approximations would be based. The variables and their associated values were as follows:

1. grade in school (we used fourth, sixth, and eighth grade students)
2. country of birth (students were classified in terms of whether they had been born in the United States or in Latin America)
3. amount of formal education (students were classified in terms of the actual number of years of formal education they had received both in the United States and in Latin America)
4. father's level of education (students were classified in terms of the actual number of years of formal education their fathers had received)
5. time in the United States (students were classified in terms of the actual number of years they had lived in the United States)

6. English proficiency (Each student was placed in one of seven categories based jointly on scores from the tests used in the participating districts for purposes of program classification and teacher judgment. Tests varied among the Language Assessment Battery, the Bilingual Syntax Measure, and the Language Assessment Scales.)

7. Spanish language proficiency (Each student was placed in one of seven categories based jointly on tests for language dominance and teacher judgment.)

Data for each of these variables were coded on cards like the facsimile presented in Exhibit 2. Once the cards were filled out, they were separated by grade level. Then, working one grade level at a time, the cards were placed two at a time in two stacks corresponding to the two training groups. Each pair of cards was matched on the seven variables as closely as possible. The bilingual resource teachers assisted in this so that approximate equality between groups would be obtained. In all cases, these teachers were satisfied that there was no systematic bias favoring one group or the other.

Following this step, cards in each of the two piles were distributed over four piles corresponding to the four testing conditions. This was done by (1) identifying sets of four students who were as nearly alike as possible and (2) assigning them randomly to testing conditions. Once again, the resource teachers were satisfied that the groups were as nearly equal as they could be made based on the available information. The numbers of students assigned to each condition is presented in Exhibit 3. Note that there were no fourth grade students in the "no testwiseness/oral administration" groups.

Assignment of Students to Conditions

Once students had been assigned to groups such that the groups were as nearly equal as possible, the groups were assigned randomly to treatments. A total of 88 students was in the group receiving Level 1 of the test. A total of 73 students was in the group receiving Level 3.

Training in Testwiseness

Students who had been selected to receive training in testwiseness received their training no more than two weeks prior to taking the CTBS. The training was conducted jointly by the investigators and by personnel from the participating districts. In this section, the training procedures, training schedules, training environments and other training factors are described.

EXHIBIT 2

STUDENT DATA CARD

Name_____ District_____

School_____ Grade_____

Country of birth_____ Years of education_____

Years of father's education_____ Time in U.S._____

English proficiency_____ Spanish proficiency_____

CTBS English %ile

Reading_____

Math_____

CTBS Español %ile

Reading_____

Math_____

Assignment to Treatment

Language of the test

English _____

Spanish _____

Testwiseness training

Yes _____

No _____

Administration

Standard _____

Oral _____

EXHIBIT 3

DISTRIBUTION OF PARTICIPANTS OVER CONDITIONS

CTBS Level 1, Grade 4

	Testwiseness Training n = 66		No Testwiseness Training n = 22	
	Oral Administration	Standard Administration	Oral Administration	Standard Administration
Spanish version	n = 20	n = 20	n = 0	n = 11
English version	n = 13	n = 13	n = 0	n = 11

CTBS Level 3, Grades 6 and 8

	Testwiseness Training n = 41		Testwiseness Training n = 32	
	Oral Administration	Standard Administration	Oral Administration	Standard Administration
Spanish version	n = 11	n = 8	n = 6	n = 11
English version	n = 11	n = 11	n = 4	n = 11

Training the Trainers

Three persons in addition to the principal investigator were scheduled to conduct the training. The three persons each came from a different district, one from District A, one from District C, and one from District D. The principal investigator conducted the training in District B. No students from District E received training in testwiseness.

Since the principal investigator had been (1) the translator of the testwiseness materials, (2) the tryout teacher, and (3) the one who actually wrote the revisions, it was natural that he be the trainer of the three persons who would conduct testwiseness training.

The three trainers plus other representatives of districts A,B,C, and D came together for an all-day training session. During the first two hours, the purposes and procedures of the study were explained in detail. The next three hours were dedicated to going over the materials and procedures for conducting testwiseness training. The majority of the participants were Spanish/English bilinguals. Accordingly, their advice was sought regarding the accuracy of the translations. An attempt was made to rid the materials of vocabulary that would be strictly regional. Their suggestions were incorporated into the final revision of the Spanish version of the materials. The final hour of the day was dedicated to instructing the trainers in procedures for selecting participants and for assigning them to treatments.

By the end of the day, all of the persons attending the training day expressed enthusiasm for the project. Each felt well-prepared to go ahead with his or her part of the study.

At the request of representatives from Districts A,B, and C, the principal investigator visited the respective districts to provide additional assistance in sample selection. In District E, the principal investigator was solely responsible for sample selection.

Training Schedules

All of the participating districts wished to avoid burdening the students with testing over and above that which is normally conducted at the end of the school year. As a result, training was conducted as early as mid-May in one district and as late as mid-June in another. The unique constraints in each district forced us to accept a variety of testing schedules. In District A and D, four one-hour sessions were conducted at each of two schools. In District B, two two-hour sessions were conducted at one school. In District C, one four-hour session was conducted at each of two schools. In all cases at least four hours of instruction were given.

Conducting the Testing

The testing, like the training, was the joint responsibility of the investigators and school personnel. Testing in Districts A and C was conducted by the same persons who conducted testwiseness training. Testing in District E was conducted by the District's classroom teachers according to standard District procedures. Testing in District B was conducted by the investigators only and in District D by the investigators and the testwiseness trainer jointly.

In District B, two examiners were used. Students were pulled out of their classrooms in small groups and were taken to special rooms where distractions would be at a minimum. Each group was tested on two successive days. Groups that were tested in the morning returned to their classrooms in the afternoon and vice versa for those who were tested in the afternoon. Students were seated at separate desks so as to minimize the likelihood of copying. In the case of the students who received the oral administration condition, the tape recording was played on a single tape player with the volume turned up to a comfortable level. As far as could be determined, each student could hear the recording well. All sessions were carefully proctored.

In District C, the students were divided into four groups according to testing conditions and taken to a trailer where they would be protected from distraction. Each group received the whole test (i.e., four subtests) in one day. The four groups received the test on four successive days. The sessions were carefully proctored.

In District D, four examiners were used. Students were pulled out of their classrooms in small groups and were taken to special rooms. The students receiving the oral administration procedures were taken to a media center consisting of two adjacent rooms where listening posts were available. The tape recording was played to them through headsets with individually adjustable volume controls. The use of the listening posts dictated that students be seated at the same table. Accordingly, proctoring was tight in order to minimize the likelihood of copying. In the case of the students who received the standard administration condition, seats were sufficiently far apart to eliminate the possibility of copying. These sessions were also proctored carefully.

In District E, students were tested in their classrooms by their own teachers according to standard district procedure.

The total number of students taking the CTBS/S was 74. The total number taking the CTBS Espanol was 87.

Analyzing the Results

Changes in the Design

At the orientation and training session in which the testwiseness trainers were trained, objections to the pretest/posttest feature of the design were raised by the school district representatives. They felt that the pretest should be avoided in order to avoid over-loading the students with testing. Classroom teachers tend to feel that testing is overdone anyway. The representatives therefore felt that the teachers and students both would be more receptive to participation if we could go to a posttest-only design. The invesigators had no objections to this change since it was felt that adequate controls had been built into the sampling plan, thus increasing the likelihood that the results would be interpretable in terms of the variables that were under study. The Program Officer was consulted regarding the change to a posttest-only design. She had no objections. Accordingly, the study was conducted in two stages. During the first stage, testwiseness training was given to the students who had been selected for such training. During the second stage the students were tested using one of four conditions as follows:

1. CTBS in English/standard administration procedure
2. CTBS in Spanish/standard administration procedure
3. CTBS in English/oral administration procedure
4. CTBS in Spanish/oral administration procedure

The Analytic Approach

As described in a previous section of this report, the study was designed as an investigation of three experimental factors: receipt of testwiseness training, form of test administration, and language of test. Each of these factors was studied at two levels: testwiseness training vs. no testwiseness training, oral vs. standard test administration, and English vs. Spanish version of the CTBS. The levels of these factors were considered as fixed and the resulting design was a 2x2x2 factorial with students serving as replicates within each of the eight treatment cells.

Analysis of this design produces a maximum of seven terms corresponding to the sources of variance in an analysis of variance table. These terms represent the three main

effects (testwiseness training, form of test administration, and language of test), the three-two way interactions (testwiseness and administration), and finally, the three-way interaction term (language and testwiseness and administration). Estimation and significance testing of each of these terms is possible as long as there are replications within the design cells so that the residual variance can be separately estimated.

At Level 3 (grades 6 and 8), each of the eight experimental cells had more than two observations, thus allowing estimation and testing of the complete model (all seven variance sources). However, at Level 1, two of the cells were empty, producing a reduced model capable of estimating and testing only five terms. Since both empty cells involved the condition of oral administration and no testwiseness training, the two-way interaction terms between these two sources and the three-way interaction term could not be separately estimated. Therefore, the residual variance term for this model included these sources of variance. For the discussion of this design, the reader may consult a standard experimental textbook such as Winer (1962) or Kirk (1968).

Actual computation of the estimates and their associated significance tests was accomplished through the use of the Statistical Analysis System (SAS) procedure labeled TLM. This procedure was chosen because both the Level 1 and Level 3 designs were unbalanced (unequal cell frequencies) and this procedure would compjte the correct sum of squares for such designs. The procedure uses a general linear model approach to allow the specification and estimation of a wide range of analytic models. While this procedure estimates four different types of estimable functions, the relevant type for the current analysis is the Type IV function. This function allows for the testing of the hypotheses as if the data had been originally derived from a balanced design. More specifically, the Type IV function estimates sum of squares associated with adding each term as the last term to the linear model. In other words, it computes the sum of squares for each term holding all the other terms of the model constant (see Helwig and Council, 1979).

All computing was performed at the University of Southern California computer facility using an IBM 370/58.

The results of these analyses are reported and discussed in the following section.

The socio-linguistic item analysis referred to earlier was structurally unrelated to the 2 x 2 x 2 factorial analysis described above. In other words, there was no attempt to study the effects of the various treatment conditions on individual item response tendencies. The socio-linguistic analysis was designed to stand alone and to provide a getter understanding of how Spanish-speaking LEP students as a group respond to paired English/Spanish items.

RESULTS

The Linguistic Analysis

As was mentioned earlier, two approaches to item analysis were used. A linguist examined each of the items from a socio-linguistic standpoint. This analysis was conducted on Levels B, C, 1, 2, and 3 of both English and Spanish versions of the tests. Additionally, a statistical analysis was limited to Levels 1 and 3 since these were the only levels actually administered to subjects in the study.

In presenting the results of the analysis we shall first deal with a comparative analysis of the overall distribution of item difficulties across four different sub-populations. Following this, we shall discuss the socio-linguistic and statistical analysis in an integrated fashion.

The comparative analysis is based on item difficulties expressed as the proportion of correct responses among four distinct groups. The groups are (1) the norming sample used by the manufacturer for the CTBS/S, (2) the sample of LES students from this study who took the CTBS/S, (3) the norming sample used by the manufacturer for the CTBS Espanol, and (4) the sample of LES students from this study who took the CTBS Espanol. The relevant data are found in Tables 3-10.

Consider first the relative performance of the English norm sample (Column 1) and the Spanish study sample (Column 4). This comparison is of interest because of the fact that it involves the two primary groups for whom the two different language versions were prepared. Other things being equal, one might hypothesize that item difficulty levels would be approximately equal for each English/Spanish item pair. Examination of the tables reveals that this hypothesis holds for a significant number of the items. If we define "approximate equality" as item pairs whose associated difficulty indices are within $\pm$ .10 of one another, we find 19 out of the 40 items on the Level 1 Vocabulary subtest that can be said to be of approximately equal difficulty. On the Level 1 Reading Comprehension subtest, 32 out of 45 items were of approximately equal difficulty. On the Level 1 Mathematics Computation subtest, 41 out of 48 items were of approximately equal difficulty. On the Level 1 Mathematics Concepts and Applications subtest, 22 out of 50 items were of approximately equal difficulty. Similar proportions of equally-difficult items were found on the Level 3 subtests. Of course, the remaining English/Spanish item pairs differed by more than $\pm$ .10. The differences were occasionally as great as $\pm$ .60. In all but a few cases, differences larger than $\pm$ .10 favored the English language sample. In other words, if an item pair exceeded the $\pm$ .10 difference in difficulty, the item was easier for the English language sample than for the Spanish

TABLE 3 : Item Difficulty in Proportion of Correct Response for Four Reference Groups: Vocabulary Subtest

Item	CTBS/S, Level 1		CTBS Español, Level 1	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.94	.71	.89	.95
2	.88	.36	.78	.97
3	.90	.57	.52	.90
4	.86	.50	.39	.46
5	.85	.60	.47	.62
6	.86	.50	.49	.78
7	.77	.22	.38	.26
8	.80	.50	.39	.56
9	.80	.67	.64	.82
10	.84	.63	.39	.54
11	.78	.48	.24	.16
12	.71	.44	.47	.71
13	.64	.39	.56	.79
14	.80	.50	.38	.55
15	.83	.61	.46	.87
16	.62	.50	.54	.85
17	.75	.80	.74	.85
18	.70	.59	.55	.62
19	.67	.42	.66	.74
20	.63	.32	.51	.72
21	.77	.39	.53	.87
22	.68	.17	.46	.64
23	.62	.29	.31	.18
24	.70	.54	.54	.77
25	.61	.26	.35	.77
26	.63	.43	.46	.76
27	.56	.32	.36	.66
28	.53	.18	.51	.66
29	.51	.48	.26	.32
30	.51	.05	.25	.18
31	.58	.50	.39	.82
32	.74	.33	.35	.44
33	.49	.20	.26	.13
34	.45	.33	.38	.59
35	.52	.00	.43	.45
36	.49	.36	.49	.76
37	.53	.19	.23	.49
38	.54	.27	.26	.14
39	.50	.24	.34	.42
40	.63	.38	.29	.54

TABLE 4 : Item Difficulty in Proportion of Correct Response for Four Reference Groups: Reading Comprehension Subtest

Item	CTBS/S, Level 1		CTBS Espanol, Level 1	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.71	.56	.59	.76
2	.86	.52	.67	.79
3	.66	.32	.45	.76
4	.79	.44	.55	.79
5	.82	.64	.64	.92
6	.82	.72	.58	.73
7	.87	.58	.58	.73
8	.73	.79	.59	.62
9	.53	.18	.44	.38
10	.78	.33	.46	.73
11	.40	.13	.42	.29
12	.67	.46	.38	.47
13	.61	.35	.40	.67
14	.76	.29	.41	.45
15	.63	.17	.31	.38
16	.82	.65	.51	.68
17	.79	.25	.53	.68
18	.74	.18	.49	.43
19	.65	.26	.39	.35
20	.48	.13	.32	.29
21	.76	.36	.43	.63
22	.68	.42	.37	.51
23	.66	.38	.47	.69
24	.63	.42	.41	.64
25	.61	.21	.37	.56
26	.75	.40	.48	.72
27	.73	.56	.40	.72
28	.78	.52	.52	.80
29	.76	.35	.48	.71
30	.67	.29	.44	.69
31	.78	.48	.47	.66
32	.62	.28	.30	.45
33	.64	.50	.44	.63
34	.68	.67	.36	.66
35	.79	.38	.51	.71
36	.75	.57	.42	.49
37	.71	.71	.49	.74
38	.68	.39	.45	.60
39	.47	.29	.26	.24
40	.52	.17	.34	.54
41	.77	.74	.48	.73
42	.60	.32	.38	.46
43	.63	.33	.32	.51
44	.60	.48	.23	.24
45	.52	.39	.36	.63

TABLE 5 : Item Difficulty in Proportion of Correct Response for Four Reference Groups: Mathematics Computation Subtest

Item	CTBS/S, Level 1		CTBS Espanol, Level 1	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.94	1.00	.95	1.00
2	.95	.96	.97	1.00
3	.91	.92	.93	1.00
4	.88	.84	.91	.87
5	.85	.88	.87	.92
6	.80	.75	.85	.92
7	.81	.76	.85	.87
8	.89	.88	.90	.97
9	.84	.75	.87	.92
10	.77	.84	.81	.78
11	.82	.68	.83	.92
12	.73	.54	.72	.75
13	.88	.75	.90	.97
14	.80	.56	.83	.84
15	.77	.68	.75	.89
16	.84	.92	.86	.97
17	.81	.88	.85	.89
18	.78	.76	.82	.87
19	.79	.72	.78	.89
20	.70	.45	.69	.67
21	.83	.96	.82	.95
22	.80	.83	.78	.76
23	.68	.48	.70	.68
24	.65	.52	.63	.64
25	.89	1.00	.92	.95
26	.92	1.00	.92	1.00
27	.87	.92	.91	.92
28	.88	.88	.88	.95
29	.81	.84	.81	.89
30	.73	.46	.77	.78
31	.75	.63	.80	.78
32	.82	.96	.85	.89
33	.76	.58	.79	.70
34	.84	.92	.86	.92
35	.73	.71	.75	.76
36	.69	.63	.75	.76
37	.85	.92	.87	.89
38	.86	.83	.87	.95
39	.70	.76	.76	.84
40	.87	.92	.87	.92
41	.83	.88	.87	.86
42	.83	.88	.86	.86
43	.84	.88	.86	.86
44	.82	.78	.84	.84
45	.80	.83	.81	.84
46	.63	.38	.70	.38
47	.65	.26	.67	.43
48	.61	.42	.69	.55

TABLE 6 : Item Difficulty in Proportion of Correct Response for Four Reference Groups: Mathematics Concepts and Applications Subtest

Item	CTBS/S, Level 1		CTBS Espanol, Level 1	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.55	.29	.28	.24
2	.85	.84	.77	.89
3	.90	.88	.81	.89
4	.79	.76	.71	.84
5	.90	.88	.84	.97
6	.85	.64	.72	.87
7	.88	.60	.73	.78
8	.71	.56	.58	.79
9	.90	.88	.79	.92
10	.83	.80	.75	.76
11	.84	.88	.79	.95
12	.82	.63	.52	.81
13	.74	.45	.59	.53
14	.51	.39	.48	.39
15	.63	.43	.49	.29
16	.76	.48	.58	.71
17	.78	.48	.66	.73
18	.89	.92	.78	.97
19	.74	.35	.62	.50
20	.86	.79	.74	.92
21	.70	.44	.57	.82
22	.66	.36	.54	.56
23	.66	.17	.48	.46
24	.60	.14	.35	.43
25	.67	.50	.60	.63
26	.81	.76	.68	.76
27	.80	.54	.63	.82
28	.79	.54	.55	.45
29	.67	.17	.54	.65
30	.72	.57	.55	.53
31	.85	.59	.59	.61
32	.75	.30	.52	.62
33	.55	.33	.37	.35
34	.56	.15	.38	.35
35	.79	.35	.63	.57
36	.55	.13	.37	.35
37	.79	.50	.70	.63
38	.72	.43	.52	.47
39	.71	.18	.56	.47
40	.60	.47	.47	.65
41	.78	.39	.60	.70
42	.80	.48	.63	.58
43	.69	.59	.47	.43
44	.63	.46	.54	.50
45	.70	.35	.59	.58
46	.65	.17	.46	.45
47	.70	.65	.56	.63
48	.63	.22	.43	.32
49	.61	.43	.60	.76
50	.66	.41	.57	.41

TABLE 7 : Item Difficulty in Proportion of Correct Response for Four Reference Groups: Vocabulary Subtest

Item	CTBS/S, Level 3		CTBS Espanol, Level 3	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.88	.43	.82	.67
2	.77	.43	.62	.46
3	.80	.43	.70	.38
4	.76	.39	.71	.58
5	.85	.52	.78	.79
6	.79	.39	.81	1.00
7	.83	.22	.49	.75
8	.77	.30	.39	.63
9	.79	.61	.71	.92
10	.65	.22	.47	.50
11	.83	.78	.67	.75
12	.69	.39	.73	.88
13	.72	.35	.57	.75
14	.77	.61	.55	.71
15	.77	.70	.73	.75
16	.71	.65	.87	.92
17	.61	.22	.58	.96
18	.60	.43	.56	.71
19	.57	.17	.73	.75
20	.62	.35	.37	.67
21	.67	.30	.44	.63
22	.74	.35	.63	.88
23	.65	.43	.44	.21
24	.61	.35	.43	.58
25	.62	.32	.61	.00
26	.63	.41	.52	.54
27	.62	.59	.54	.71
28	.61	.18	.22	.25
29	.50	.33	.51	.63
30	.65	.29	.39	.21
31	.55	.67	.67	.75
32	.58	.38	.52	.67
33	.56	.43	.47	.54
34	.53	.57	.54	.48
35	.48	.38	.56	.59
36	.49	.45	.37	.64
37	.48	.42	.25	.36
38	.46	.21	.23	.36
39	.50	.47	.39	.68
40	.49	.21	.48	.87

TABLE 8: Item Difficulty in Proportion of Correct Response for Four Reference Groups: Reading Comprehension Subtest

Item	CTBS/S, Level 3		CTBS Espanol, Level 3	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.90	.70	.78	.88
2	.78	.35	.81	.88
3	.71	.22	.59	.75
4	.69	.39	.60	.63
5	.67	.39	.47	.17
6	.72	.59	.56	.43
7	.75	.23	.57	.50
8	.58	.43	.47	.75
9	.53	.13	.34	.50
10	.49	.17	.44	.29
11	.81	.83	.68	.92
12	.63	.48	.53	.65
13	.68	.13	.46	.63
14	.55	.39	.36	.50
15	.38	.22	.30	.37
16	.49	.48	.26	.30
17	.80	.39	.59	.63
18	.82	.70	.65	.80
19	.72	.30	.51	.60
20	.68	.35	.53	.77
21	.82	.70	.59	.73
22	.76	.43	.49	.50
23	.69	.43	.48	.47
24	.85	.87	.64	.97
25	.66	.48	.43	.10
26	.74	.61	.46	.72
27	.64	.61	.41	.45
28	.76	.48	.45	.45
29	.67	.26	.41	.45
30	.54	.09	.32	.52
31	.55	.43	.34	.59
32	.76	.61	.46	.72
33	.82	.73	.56	.79
34	.67	.18	.40	.62
35	.67	.64	.40	.69
36	.61	.23	.44	.43
37	.67	.55	.46	.50
38	.46	.36	.38	.64
39	.45	.18	.30	.52
40	.48	.14	.26	.19
41	.52	.14	.41	.46
42	.50	.23	.27	.44
43	.53	.36	.26	.26
44	.38	.18	.21	.31
45	.48	.23	.27	.27

TABLE 9: Item Difficulty in Proportion of Correct Response for Four Reference Groups: Mathematics Computation Subtest

Item	CTBS/S, Level 3		CTBS Espanol, Level 3	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.88	.87	.89	.93
2	.71	.65	.69	.60
3	.59	.41	.55	.30
4	.59	.18	.54	.30
5	.64	.43	.51	.40
6	.63	.57	.56	.40
7	.73	.78	.73	.60
8	.68	.87	.72	.70
9	.65	.43	.55	.47
10	.68	.57	.67	.47
11	.57	.52	.61	.43
12	.70	.70	.73	.67
13	.82	.91	.80	.87
14	.63	.30	.52	.40
15	.65	.43	.62	.53
16	.72	.91	.71	.63
17	.53	.17	.46	.30
18	.48	.26	.45	.23
19	.47	.17	.40	.13
20	.67	.73	.66	.73
21	.55	.36	.56	.52
22	.74	.83	.74	.73
23	.81	.87	.80	.80
24	.67	.61	.65	.57
25	.74	.83	.78	.73
26	.64	.68	.64	.69
27	.70	.83	.66	.70
28	.81	.91	.78	.83
29	.64	.52	.61	.47
30	.64	.30	.59	.30
31	.62	.23	.53	.27
32	.62	.26	.56	.30
33	.69	.73	.66	.73
34	.70	.62	.72	.70
35	.58	.63	.52	.60
36	.69	.90	.59	.67
37	.69	.90	.59	.67
38	.54	.42	.48	.31
39	.52	.56	.51	.43
40	.65	.56	.57	.50
41	.52	.47	.48	.23
42	.65	.76	.56	.77
43	.50	.35	.44	.33
44	.50	.41	.45	.36
45	.67	.76	.59	.79
46	.47	.59	.47	.25
47	.40	.24	.34	.39
48	.52	.69	.48	.70

TABLE 10: Item Difficulty in Proportion of Correct Response for Four Reference Groups: Mathematics Concepts and Applications Subtest

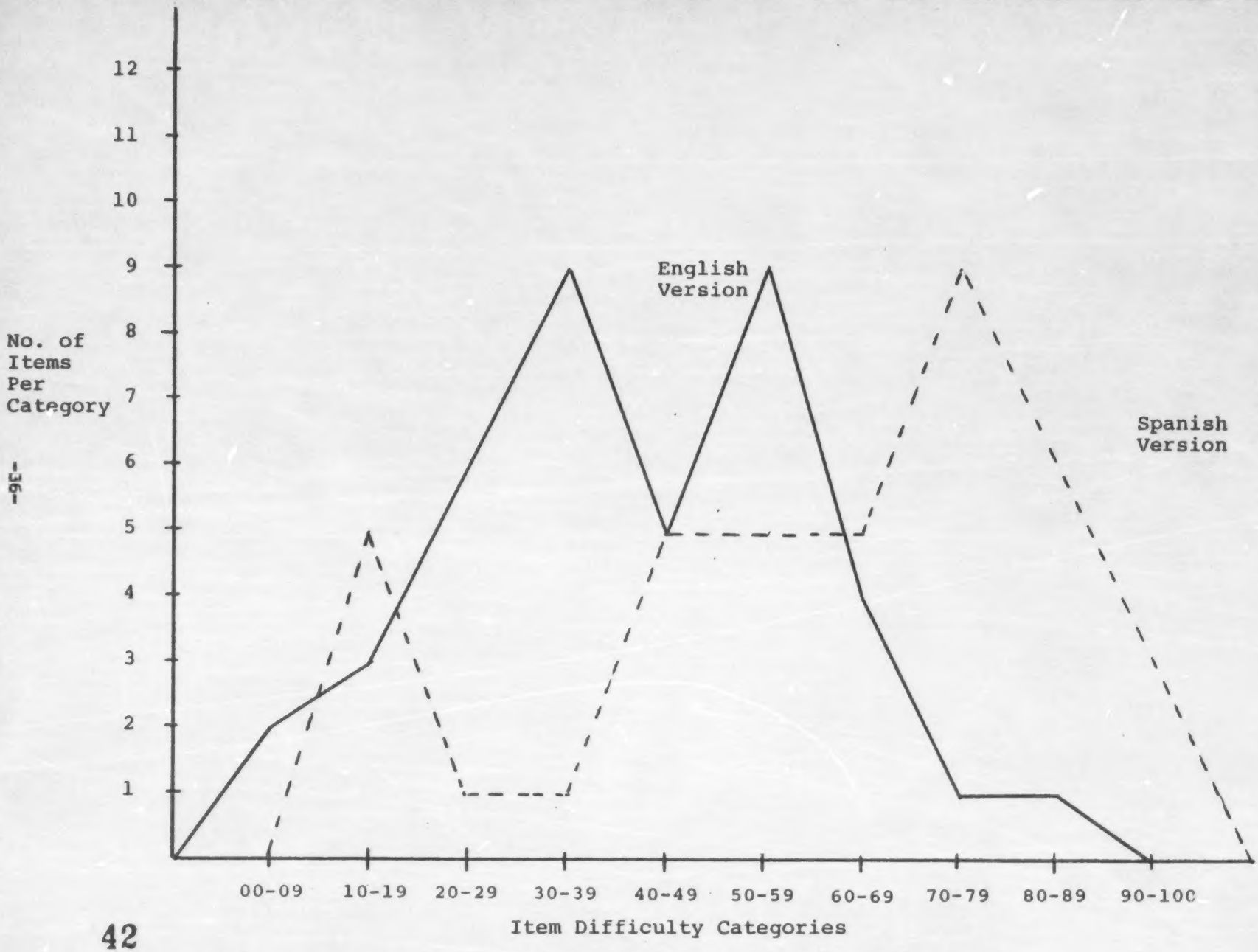
Item	CTBS/S, Level 3		CTBS Espanol, Level 3	
	Norm Sample	Study Sample	Norm Sample	Study Sample
1	.84	.95	.87	.73
2	.55	.27	.44	.23
3	.82	.64	.77	.73
4	.60	.50	.49	.40
5	.37	.45	.40	.23
6	.67	.43	.56	.47
7	.63	.25	.59	.27
8	.76	.71	.64	.50
9	.51	.27	.41	.43
10	.58	.45	.47	.43
11	.64	.36	.60	.40
12	.69	.71	.64	.57
13	.65	.67	.57	.47
14	.39	.30	.40	.27
15	.70	.33	.55	.41
16	.67	.43	.44	.47
17	.73	.67	.60	.57
18	.67	.33	.49	.40
19	.58	.48	.47	.33
20	.65	.38	.51	.70
21	.59	.52	.38	.45
22	.62	.05	.56	.50
23	.74	.73	.62	.73
24	.62	.33	.48	.28
25	.54	.33	.44	.33
26	.60	.62	.68	.57
27	.76	.82	.69	.63
28	.77	.62	.75	.50
29	.84	.90	.88	.60
30	.65	.52	.59	.23
31	.66	.33	.61	.25
32	.84	.90	.80	.77
33	.49	.30	.39	.34
34	.57	.29	.47	.21
35	.47	.10	.36	.10
36	.62	.50	.63	.63
37	.68	.62	.67	.50
38	.51	.32	.51	.47
39	.62	.11	.63	.50
40	.66	.32	.60	.57
41	.63	.30	.65	.40
42	.65	.58	.67	.64
43	.48	.21	.35	.14
44	.71	.68	.62	.45
45	.81	.83	.73	.70
46	.70	.41	.60	.62
47	.59	.22	.55	.38
48	.67	.41	.60	.37
49	.64	.56	.60	.41
50	.55	.39	.50	.30

language sample. This stimulated our curiosity and led us to examine these difficult items more carefully in an attempt to generate hypotheses about why these items would be harder for a Spanish-speaking group than for an English-speaking group. This analysis will be presented later in this section of the report.

Consider now the relative performance of the CTBS Espanol norming sample vis a vis both the CTBS/S norming sample and the CTBS Espanol study sample. This comparison is of interest since the CTBS Espanol norming sample was rated as bilingual-biliterate. Examination of Tables 3-10 reveals that the Spanish language test was considerably easier for the LEP sample than it was for the bilingual-biliterate sample. This difference, though present in all subtests, was most pronounced on the vocabulary and reading comprehension subtests. This pointed to the fact that the LEP group was considerably more proficient in the Spanish language than was the bilingual-biliterate group. Needless to say, given that the bilingual-biliterate subjects found the CTBS Espanol to be more difficult than the LEP group found it, the item by item comparison of the performance of the bilingual-biliterate group on the CTBS Espanol with the performance of the norming group on the CTBS/S favored the latter. In other words, the bilingual-biliterate group was less proficient in Spanish than was the CTBS/S norming group in English. We have no information on the performance of the bilingual-biliterate norming group on the CTBS/S.

Consider now the relative performance of one LEP group on the CTBS/S and the performance of the matching LEP group on the CTBS Espanol (see Tables 3-10, columns 2 and 4). Examination of the Tables reveals that for LEP students, the Spanish version of the test was considerably easier than the English version. This was expected. The importance of this study lies not in "finding" this difference but rather in documenting its magnitude in each of the subtests separately and in identifying specific features of the test materials and procedures that might enable us to suggest ways to improve estimates of achievement in LEP students.

For purposes of illustrating the magnitude of differences on the various subtests within levels, Tables 3-10 were interpreted in the form of frequency polygons. These are presented in Figures 1-8. Figures 1 and 5 show that the CTBS Espanol Vocabulary subtests were considerably easier than were the CTBS/S Vocabulary subtests. Figures 2 and 6 reflect a similar relation for the Reading Comprehension subtests. Figures 3 and 7 reflect no noticeable difference in difficulty for the Mathematics Computation subtests. This presumably reflects the fact that language (Spanish and English) has no effect on performance. Figures 4 and 8 reveal that there probably were differences in difficulty for Level 1 of the test but not for Level 3 for the Mathematics Concepts and Applications subtests. Tests of these differences will be presented later.



42

FIGURE 1: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL ONE VOCABULARY SUBTEST

43

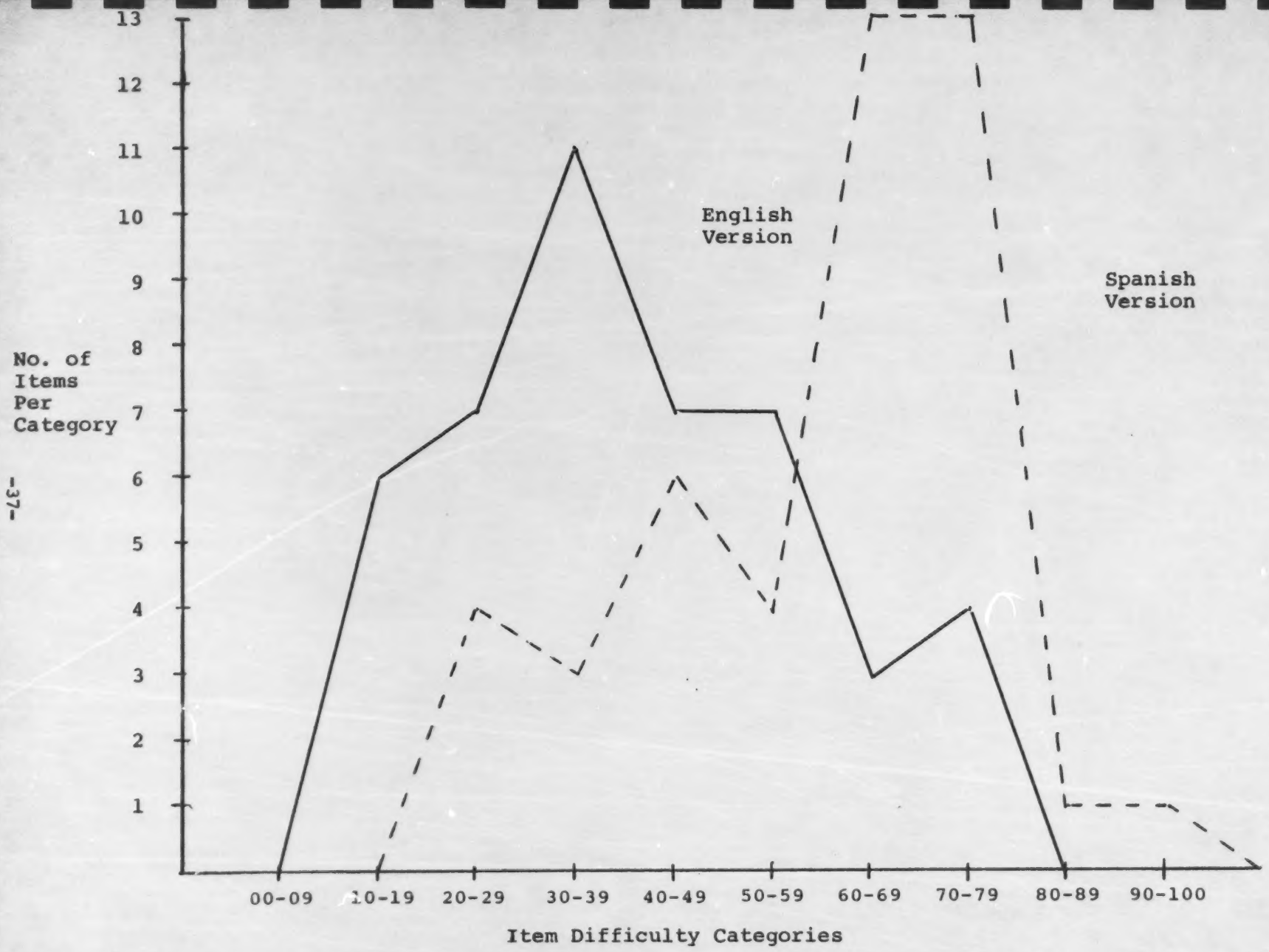


FIGURE 2: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL ONE READING COMPREHENSION SUBTEST

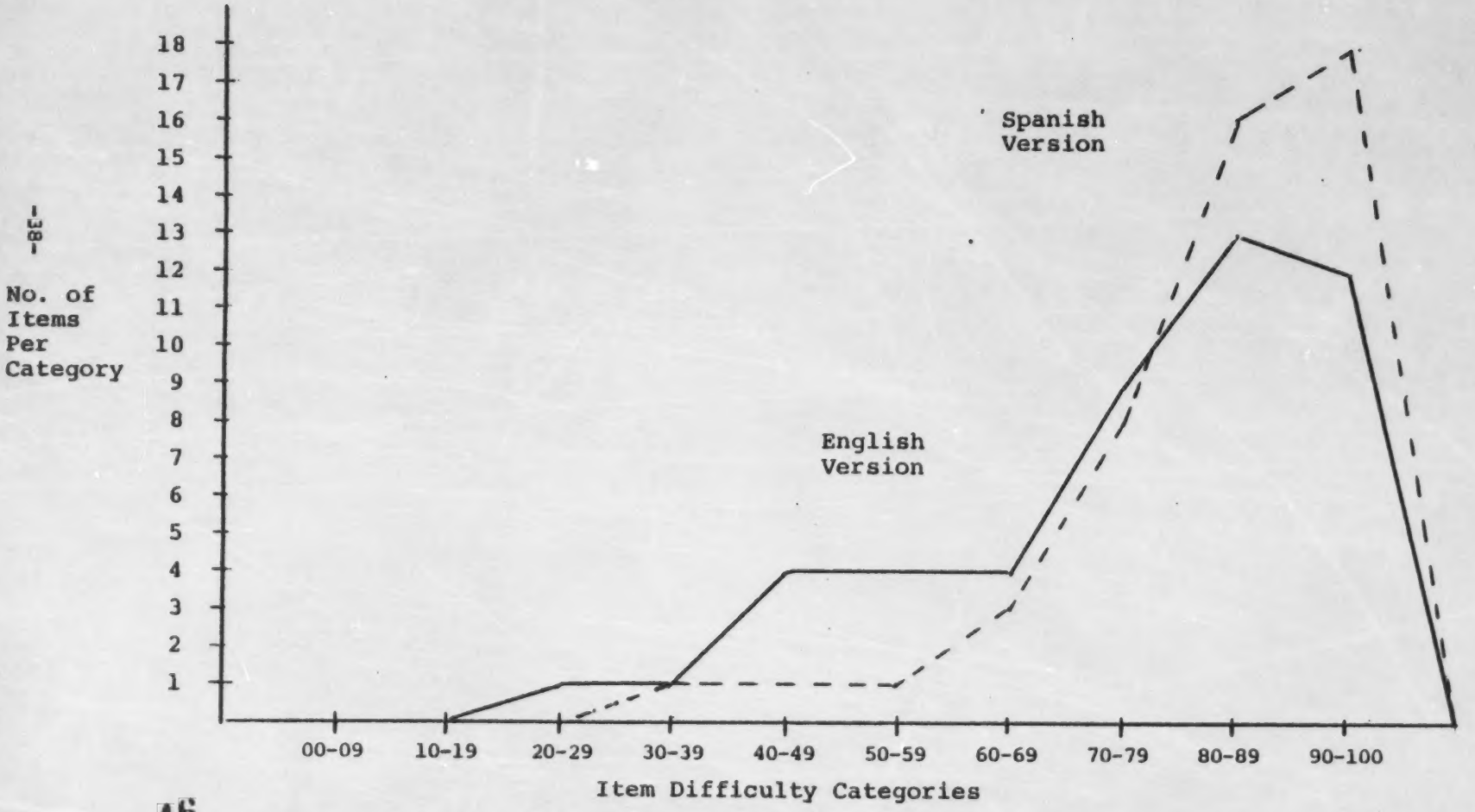


Figure 3: Distribution of Item Difficulties On the Level One Mathematics Computation Subtest

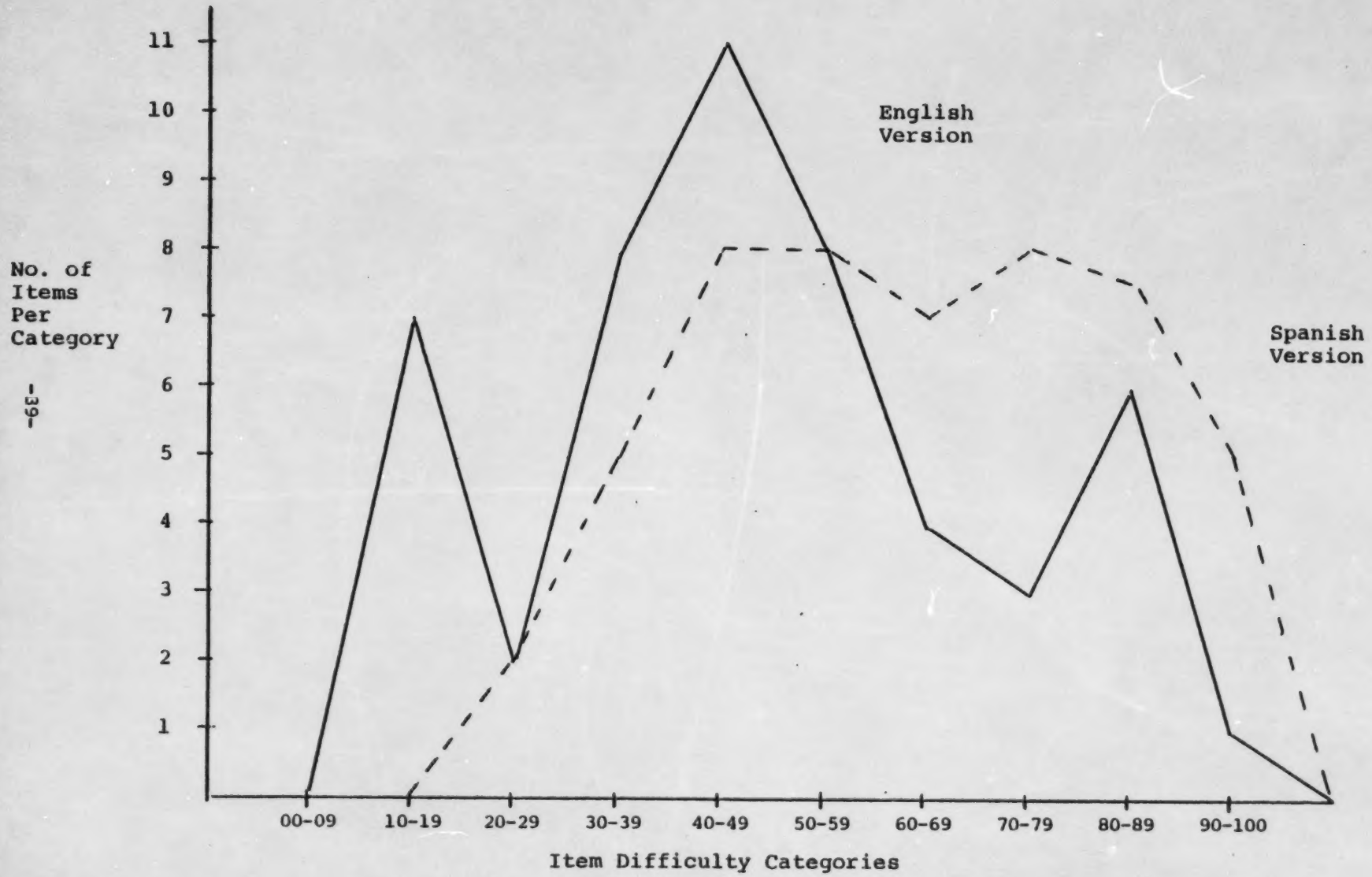


FIGURE 4: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL ONE MATH CONCEPTS AND APPLICATIONS SUBTEST

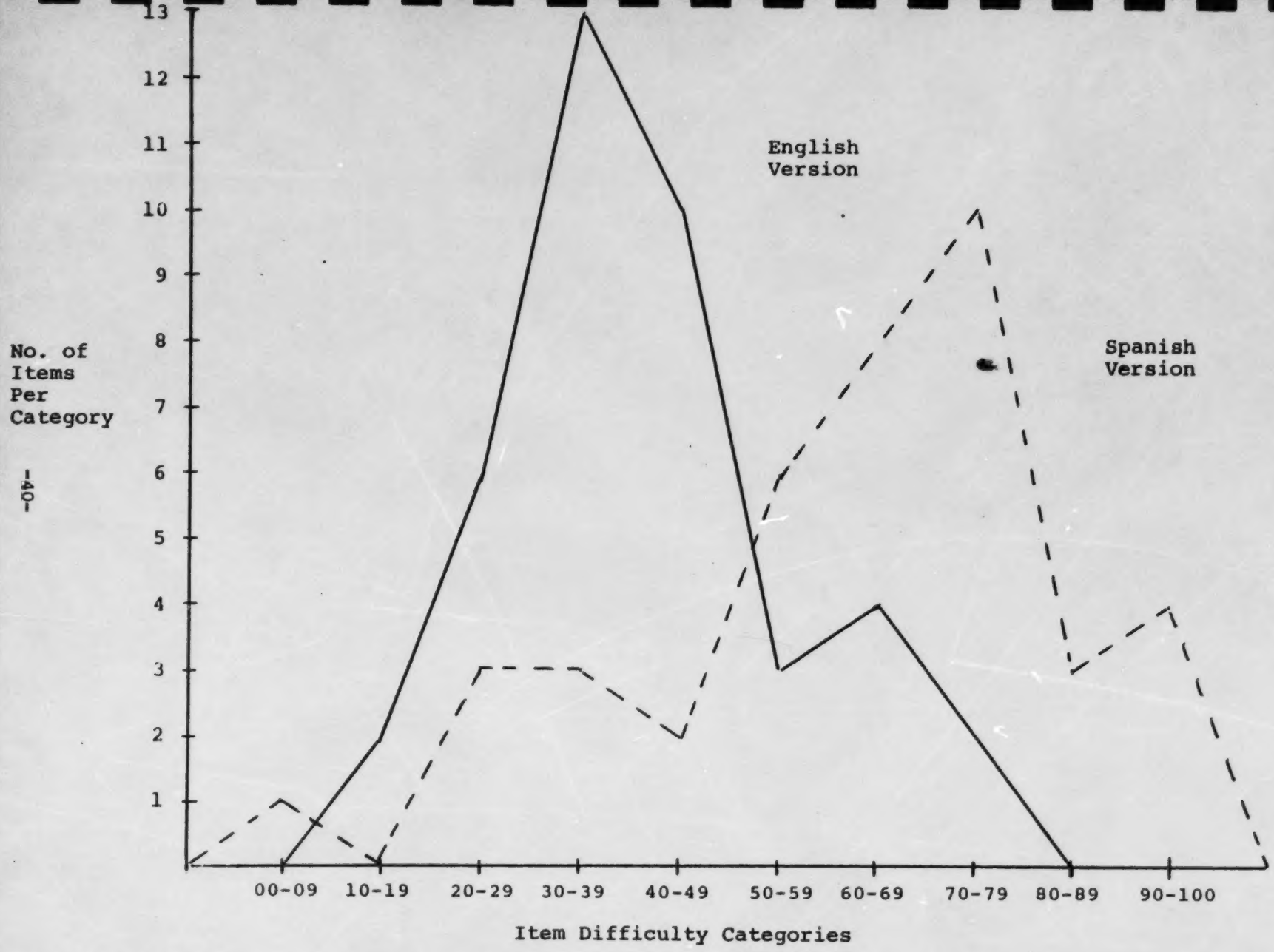


FIGURE 5: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL THREE VOCABULARY SUBTEST

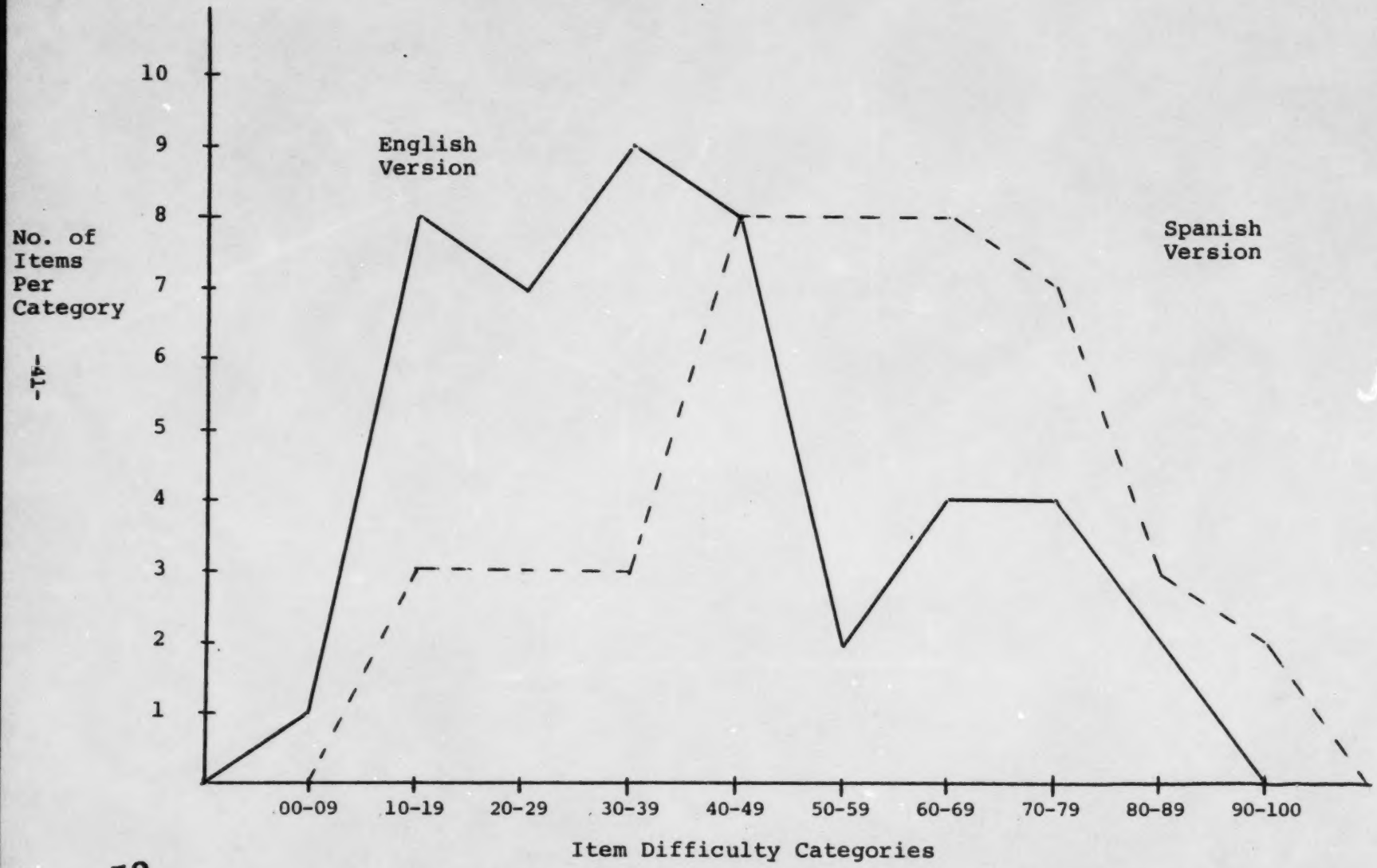


FIGURE 6: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL THREE READING COMPREHENSION SUBTEST

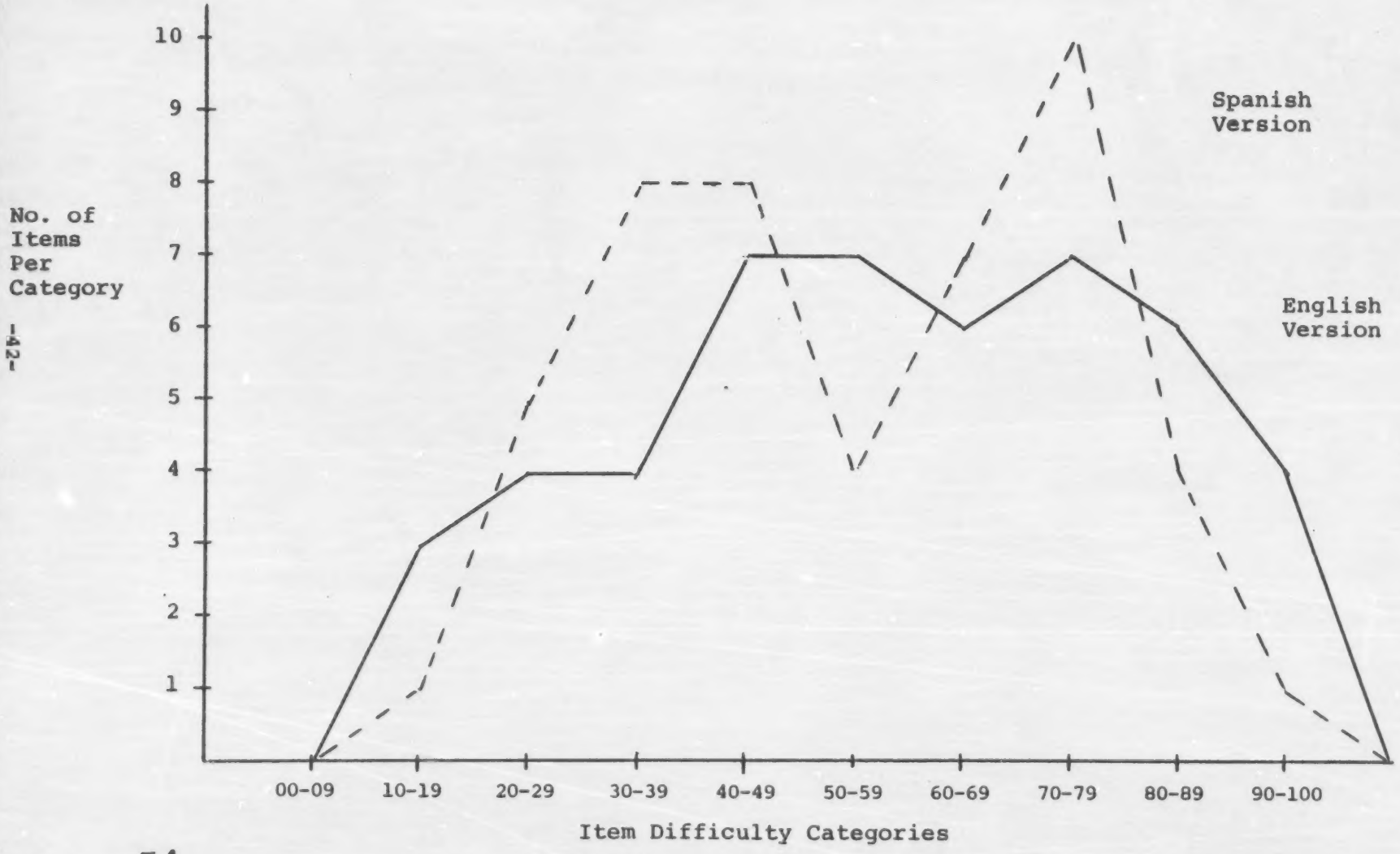


FIGURE 7: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL THREE MATHEMATICS COMPUTATION SUBTEST

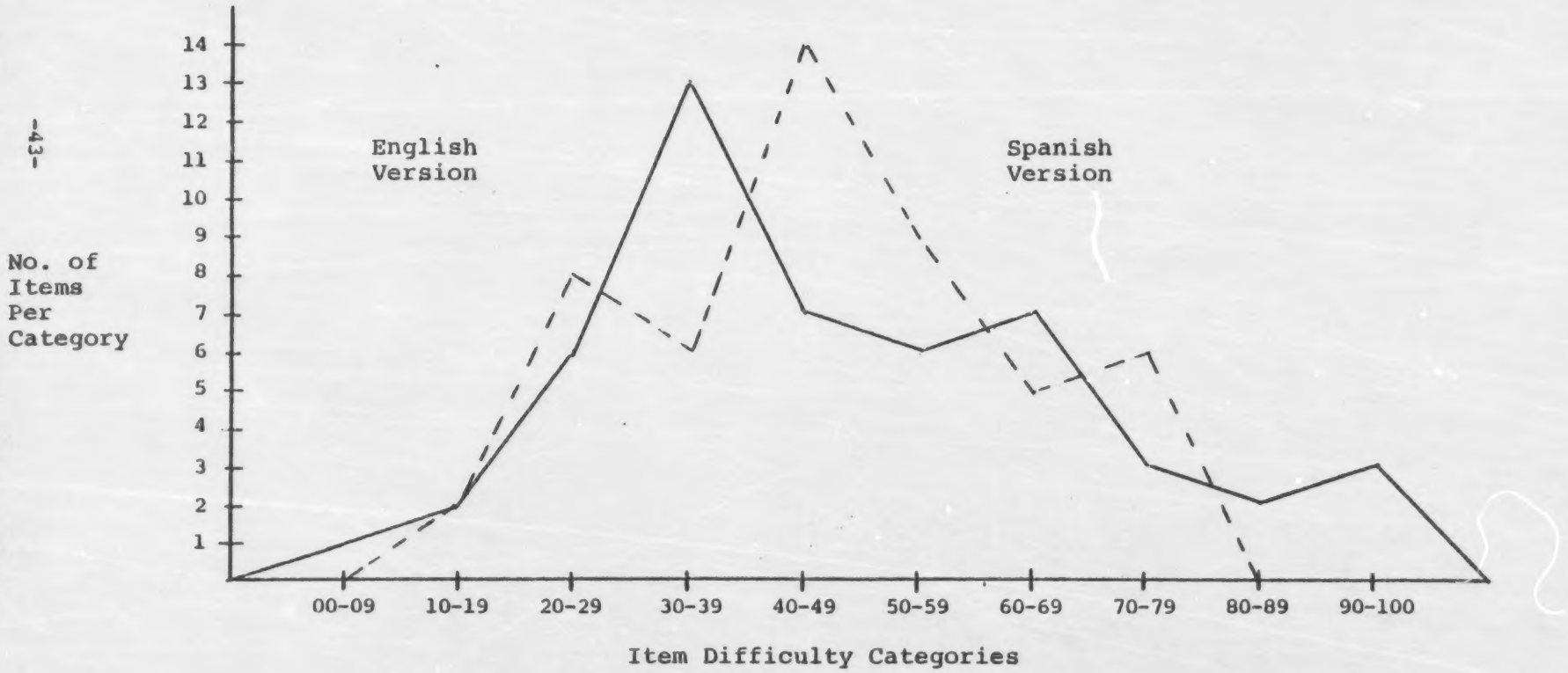


FIGURE 8: DISTRIBUTION OF ITEM DIFFICULTIES ON THE LEVEL THREE MATH CONCEPTS AND APPLICATIONS SUBTEST

To this point, we have established that:

1. the CTBS/S is easier for English-dominant students than is the CTBS Espanol for Spanish-dominant students
2. the CTBS Espanol is easier for Spanish dominant students than it is for bilingual-biliterate students
3. the CTBS/S is easier for English-dominant students than is the CTBS Espanol for bilingual-biliterate students.
4. the CTBS Espanol is easier for Spanish-dominant students than is the CTBS/S. This is more clearly the case for the vocabulary and reading comprehension subtests than it is for the mathematics computation and concepts and applications subtests. There may even be a difference in difficulty in the latter two subtests with mathematics computation being easier than mathematics concepts and applications.

The preceding analysis makes clear that LES students encounter significant difficulties with the CTBS/S and, in some cases, even with the CTBS Espanol. We may now present the results of an inquiry into the specific nature of those difficulties. In the conduct of the inquiry, Spanish/English item pairs were grouped according to their levels of relative difficulty. The item pairs of greatest interest were (1) those in which both members of the pair were easy ($\geq .70$), (2) those in which both members of the pair were difficult ($\leq .30$), and (3) those in which the Spanish member of the pair was easier than the English member (difference $\geq .10$). Having identified the relevant subsets of item pairs, the frequency of response to each of the response options was examined. These frequencies are presented as proportion of frequency in Tables 11-18. The inquiry led to a number of generalizations about response tendencies.

First, items in either language tended to be easy or hard as a function of the frequency with which the key words are used in the respective languages. Thus in the vocabulary subtest where items have the form:

> Given: a two to three word phrase where one of the words is underlined plus four single-word response options
>
> Required: select the response option that means the same as or nearly the same as the underlined word.

TABLE 11: Proportion of All Responses to Each Response Option on the Level One Vocabulary Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.04	.71*	.08	.17	.05	.95*	.00	.00
2	.23	.32	.09	.36*	.00	.00	.03	.97*
3	.57*	.10	.00	.33	.90*	.03	.03	.05
4	.17	.50*	.13	.21	.21	.46*	.13	.21
5	.60*	.12	.16	.12	.62*	.08	.16	.14
6	.18	.50*	.23	.09	.08	.78*	.03	.11
7	.39	.13	.22*	.26	.23	.10	.26*	.41
8	.09	.50*	.32	.09	.14	.57*	.14	.16
9	.17	.67*	.04	.13	.00	.82*	.05	.13
10	.17	.04	.17	.63*	.18	.15	.13	.54*
11	.48*	.13	.19	.30	.16*	.26	.11	.47
12	.40	.12.	.04	.44*	.21	.05	.03	.71*
13	.00	.39*	.30	.30	.13	.79*	.18	.00
14	.29	.50*	.08	.13	.26	.55*	.05	.13
15	.04	.13	.22	.61*	.04	.03	.05	.87*
16	.21	.13	.50*	.17	.05	.08	.85*	.03
17	.04	.08	.08	.80*	.00	.03	.13	.85*
18	.09	.18	.59*	.14	.16	.08	.62*	.14
19	.29	.42*	.17	.13	.08	.74*	.11	.08
20	.40	.24	.04	.32*	.26	.03	.00	.72*
21	.13	.04	.43	.39*	.00	.08	.05	.87*
22	.13	.17*	.63	.08	.05	.64*	.18	.13
23	.25	.13	.29*	.33	.26	.05	.18*	.50
24	.21	.54	.08	.17	.18	.77*	.05	.00
25	.09	.39	.26	.26*	.03	.15	.05	.77*
26	.43*	.13	.39	.04	.76*	.16	.05	.03
27	.45	.32*	.14.	.09	.16	.66*	.13	.05
28	.27	.41	.18*	.14	.03	.26	.66*	.05
29	.10	.10	.48*	.33	.03	.11	.54*	.32
30	.64	.14	.05*	.18	.62	.05	.18*	.15
31	.32	.14	.05	.50*	.02	.05	.05	.82*
32	.33	.29	.05	.33*	.46	.00	.10	.44*
33	.25	.35	.20*	.20	.21	.28	.13*	.38
34	.14	.29	.33*	.24	.11	.08	.59*	.22
35	.24	.33	.00*	.43	.24	.16	.45*	.16
36	.36*	.27	.23	.14	.76*	.08	.03	.14
37	.19*	.48	.19	.14	.49*	.14	.11	.26
38	.27	.36	.09	.27*	.17	.54	.14	.14*
39	.52	.24	.24*	.00	.17	.28	.42*	.14
40	.05	.43	.14	.38*	.20	.11	.14	.54*

*Indicates correct answer

TABLE 12: Proportion of All Responses to Each Response Option on the Level One Reading Comprehension Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.16	.08	.20	.56*	.08	.00	.16	.76*
2	.52*	.13	.04	.30	.79*	.00	.08	.13
3	.32	.20	.16	.32*	.08	.08	.08	.76*
4	.32	.44*	.12	.12	.18	.79*	.00	.03
5	.08	.12	.16	.64*	.00	.05	.03	.92*
6	.20	.72*	.04	.04	.14	.73*	.05	.08
7	.13	.17	.13	.58*	.00	.19	.08	.73*
8	.13	.79*	.04	.04	.32	.62*	.03	.03
9	.18	.36	.18*	.27	.11	.27	.38*	.24
10	.00	.33*	.21	.46	.00	.73*	.11	.16
11	.13*	.46*	.08	.25	.13	.47*	.05	.34
12	.21	.46*	.08	.25	.13	.47*	.05	.34
13	.26	.22	.17	.35*	.06	.08	.19	.67*
14	.29*	.17	.25	.29	.45*	.34	.03	.18
15	.29	.13	.42	.17*	.03	.05	.54	.38*
16	.22	.09	.65*	.04	.13	.05	.68*	.13
17	.21	.25*	.54	.00	.16	.68*	.16	.00
18	.36	.18*	.36	.09	.27	.43*	.22	.08
19	.09	.30	.35	.26*	.11	.30	.24	.35*
20	.13*	.22	.30	.35	.29*	.05	.21	.45
21	.36*	.08	.36	.20	.63*	.18	.08	.11
22	.17	.21	.21	.42*	.14	.23	.11	.51*
23	.08	.13	.38*	.42	.03	.06	.69*	.22
24	.08	.42*	.33	.17	.08	.64*	.08	.19
25	.21*	.21	.42	.17	.56*	.08	.17	.19
26	.20	.40*	.36	.04	.11	.72*	.08	.08
27	.00	.56*	.20	.24	.00	.72*	.14	.14
28	.04	.52*	.44	.00	.00	.80*	.20	.00
29	.12	.12	.40	.36*	.11	.06	.11	.71*
30	.29*	.25	.33	.13	.69*	.20	.03	.09
31	.16	.28	.08	.48*	.09	.20	.06	.66*
32	.20	.28*	.44	.08	.00	.45*	.42	.12
33	.08	.17	.25	.50*	.03	.23	.11	.63*
34	.04	.67*	.17	.13	.03	.66*	.11	.20
35	.04	.38	.38*	.21	.03	.20	.71*	.06
36	.17	.17	.09	.57*	.23	.17	.11	.49*
37	.17	.04	.71*	.08	.03	.06	.74*	.17
38	.09	.39*	.35	.17	.17	.60*	.11	.11
39	.29*	.21	.21	.29	.24*	.41	.18	.18
40	.17*	.17	.22	.43	.54*	.14	.23	.09
41	.04	.13	.74*	.09	.06	.12	.73*	.09
42	.12	.28	.28	.32*	.26	.17	.11	.46*
43	.13	.33	.21	.33*	.14	.26	.09	.51*
44	.13	.17	.22	.48*	.29	.18	.29	.24*
45	.26	.26	.09	.39*	.23	.06	.09	.63*

*Indicates correct answer

TABLE 13: Proportion of All Response to Each Response Option on the Level One Mathematics Computation Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.00	1.00*	.00	.00	.00	1.00*	.00	.00
2	.04	.96*	.00	.00	.00	1.00*	.00	.00
3	.00	.92*	.00	.00	.00	1.00*	.00	.00
4	.16	.00	.00	.84*	.11	.03	.00	.87*
5	.08	.88*	.00	.04	.08	.92*	.00	.00
6	.00	.21	.75*	.04	.00	.08	.92*	.00
7	.04	.12	.08	.76*	.00	.13	.00	.87*
8	.04	.04	.88*	.04	.00	.03	.97*	.00
9	.04	.13	.75*	.08	.03	.00	.92*	.05
10	.04	.04	.83*	.08	.14	.18	.78*	.00
11	.68*	.20	.00	.12	.92*	.03	.03	.03
12	.00	.08	.38	.54*	.06	.06	.14	.75*
13	.08	.75*	.13	.04	.03	.97*	.00	.00
14	.56*	.00	.28	.16	.84*	.00	.11	.05
15	.00	.32	.00	.68*	.00	.11	.00	.89*
16	.00	.08	.92*	.00	.00	.00	.97*	.03
17	.08	.88*	.04	.00	.05	.89*	.00	.05
18	.12	.76*	.08	.04	.03	.87*	.05	.05
19	.04	.72*	.12	.12	.03	.89*	.00	.08
20	.27	.05	.23	.45*	.11	.08	.14	.67*
21	.00	.04	.96*	.00	.00	.03	.95*	.03
22	.17	.00	.83*	.00	.18	.03	.76*	.03
23	.13	.17	.22	.48*	.05	.11	.16	.68*
24	.00	.17	.52*	.30	.08	.17	.64*	.11
25	.00	1.00*	.00	.00	.05	.95*	.00	.00
26	.00	1.00*	.00	.00	.00	1.00*	.00	.00
27	.92*	.08	.00	.00	.92*	.08	.00	.00
28	.18	.00	.88*	.04	.05	.00	.95*	.00
29	.00	.08	.08	.83*	.03	.08	.00	.89*
30	.46*	.04	.08	.42	.78*	.05	.00	.16
31	.04	.25	.63*	.08	.03	.14	.78*	.05
32	.00	.04	.96*	.00	.03	.08	.89*	.00
33	.17	.58*	.08	.17	.16	.70*	.03	.11
34	.08	.00	.92*	.00	.08	.00	.92*	.00
35	.04	.21	.04	.71*	.05	.16	.03	.76*
36	.04	.08	.25	.63*	.08	.05	.11	.76*
37	.00	.00	.92*	.08	.00	.03	.89*	.08
38	.08	.00	.83*	.08	.03	.00	.95*	.03
39	.24	.76*	.00	.00	.16	.84*	.00	.00
40	.92*	.00	.04	.04	.92*	.03	.00	.05
41	.08	.88*	.04	.00	.05	.86*	.05	.03
42	.88*	.00	.00	.12	.86*	.03	.00	.10
43	.88*	.04	.04	.04	.86*	.08	.00	.05
44	.00	.78*	.04	.17	.00	.84*	.08	.08
45	.83	.00	.08	.08	.84*	.05	.08	.03
46	.13	.17	.33	.38*	.14	.22	.27	.38*
47	.26*	.17	.22	.35	.43*	.14	.08	.35
48	.13	.42*	.25	.21	.11	.55*	.18	.16

*Indicates correct answer

TABLE 14: Proportion of all Responses to Each Response Option on the Level One Mathematics Concepts and Applications Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.29*	.24	.24	.24	.24*	.30	.22	.24
2	.04	.12	.00	.84*	.08	.03	.00	.89*
3	.00	.00	.88*	.13	.08	.00	.89*	.03
4	.16	.00	.76*	.08	.11	.05	.84*	.00
5	.08	.88*	.00	.04	.00	.97*	.03	.00
6	.16	.64*	.08	.12	.05	.87*	.03	.05
7	.60*	.20	.16	.04	.78*	.11	.03	.08
8	.56*	.24	.08	.12	.79*	.03	.05	.13
9	.00	.04	.88*	.08	.03	.05	.92*	.00
10	.04	.12	.80	.04	.11	.11	.76*	.03
11	.00	.88	.12	.00	.00	.95*	.05	.00
12	.63*	.08	.08	.21	.81	.05	.05	.08
13	.18	.45*	.05	.32	.08	.53*	.00	.39
14	.13	.04	.43	.39*	.11	.03	.47	.39*
15	.13	.09	.43*	.35	.08	.05	.29*	.58
16	.16	.48*	.32	.04	.00	.71*	.24	.05
17	.16	.12	.48*	.24	.03	.24	.73*	.00
18	.04	.04	.00	.92*	.00	.00	.03	.97*
19	.09	.35*	.17	.39	.00	.50*	.13	.29
20	.00	.79*	.17	.04	.00	.92*	.03	.06
21	.00	.44*	.40	.16	.00	.82*	.16	.03
22	.24	.12	.36*	.28	.08	.06	.56*	.31
23	.17*	.21	.25	.38	.46*	.24	.08	.22
24	.14*	.33	.33	.19	.43*	.14	.26	.17
25	.25	.50*	.17	.08	.11	.63*	.05	.21
26	.12	.12	.76*	.00	.15	.15	.76*	.13
27	.54*	.17	.08	.21	.82*	.03	.05	.11
28	.54*	.08	.29	.08	.45*	.08	.45	.03
29	.08	.17*	.67	.08	.11	.65*	.14	.11
30	.22	.57*	.04	.17	.18	.53*	.11	.18
31	.05	.36	.59*	.00	.13	.26	.61*	.11
32	.30*	.30	.09	.30	.62*	.14	.14	.11
33	.38	.00	.33*	.29	.38	.11	.35*	.16
34	.50	.15*	.15	.20	.51	.35*	.08	.05
35	.05	.35	.20	.40	.06	.57*	.20	.17
36	.26	.22	.13*	.39	.05	.32	.35*	.27
37	.00	.41	.50*	.09	.05	.16	.63*	.16
38	.43*	.22	.09	.26	.47*	.19	.06	.28
39	.09	.18*	.55	.18	.08	.47*	.25	.19
40	.26	.47*	.16	.11	.06	.65*	.24	.06
41	.39*	.43	.17	.00	.70*	.03	.16	.11
42	.04	.48*	.39	.09	.05	.58*	.21	.16
43	.05	.59*	.14	.23	.16	.43*	.19	.22
44	.08	.46*	.17	.29	.08	.50*	.24	.18
45	.09	.43	.35*	.13	.11	.03	.58*	.28
46	.35	.17*	.13	.35	.26	.45*	.03	.26
47	.13	.65*	.13	.09	.18	.63*	.11	.08
48	.43	.35	.22*	.00	.27	.38	.32*	.03
49	.38	.43*	.14	.05	.14	.76*	.03	.08
50	.42*	.25	.13	.21	.41*	.24	.19	.16

*Indicates correct answer

TABLE 15: Proportion of All Responses to Each Response Option on the Level Three Vocabulary Subtest

	English				Spanish			
Item	1	2	3	4	1	2	3	4
1	.43*	.17	.04	.35	.67*	.13	.00	.21
2	.35	.43*	.09	.13	.29	.46*	.08	.17
3	.26	.43*	.26	.04	.25	.38*	.33	.04
4	.39*	.22	.13	.26	.58*	.08	.04	.29
5	.26	.13	.52*	.09	.13	.08	.79*	.00
6	.26	.39*	.22	.13	.00	1.00*	.00	.00
7	.22	.17	.39	.22*	.13	.00	.13	.75*
8	.26	.30*	.13	.30	.13	.63*	.04	.21
9	.13	.61*	.04	.22	.04	.92*	.00	.04
10	.30	.43	.04	.22*	.13	.33	.04	.50*
11	.13	.78*	.09	.00	.08	.75*	.13	.04
12	.13	.13	.35	.39*	.04	.04	.04	.88*
13	.35	.22	.35*	.09	.13	.04	.75*	.08
14	.04	.30	.61*	.04	.21	.08	.71*	.00
15	.80*	.04	.04	.22	.75*	.13	.04	.08
16	.04	.17	.13	.65*	.00	.08	.00	.92*
17	.22*	.43	.09	.26	.96*	.00	.00	.04
18	.30	.00	.43*	.26	.96*	.00	.00	.04
19	.65	.09	.09	.17*	.17	.08	.00	.75*
20	.13	.43	.35*	.09	.08	.17	.67*	.08
21	.30*	.26	.43	.00	.63*	.17	.17	.04
22	.35	.26	.09	.30	.88*	.08	.04	.00
23	.30	.17	.43*	.09	.38	.42	.21	.00
24	.35*	.35	.26	.04	.58*	.29	.13	.00
25	.27	.18	.32*	.23	.17	.79	.00*	.04
26	.32	.14	.14	.41*	.33	.13	.00	.54*
27	.59*	.09	.14	.18	.71*	.17	.08	.04
28	.32	.18	.23	.27	.42	.25*	.13	.21
29	.14	.14	.38	.33*	.04	.08	.25	.63
30	.62	.10	.29*	.00	.50	.17	.21	.13
31	.14	.67*	.05	.14	.13	.75*	.04	.08
32	.15	.33	.38*	.24	.08	.08	.67*	.17
33	.38	.10	.10	.43	.29	.13	.04	.54*
34	.56*	.10	.24	.10	.48*	.04	.22	.26
35	.14	.38	.10	.38	.09	.32	.00	.59
36	.25	.10	.20	.45	.14	.09	.14	.64
37	.11	.42*	.37	.11	.14	.36*	.41	.09
38	.16	.37	.26	.21*	.23	.27	.14	.36*
39	.47*	.21	.05	.26	.68*	.18	.09	.05
40	.32	.21*	.32	.16	.04	.87*	.00	.09

*Indicates correct answer

TABLE 16: Proportion of All Responses to Each Response Option on the Level Three Reading Comprehension Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.70*	.13	.09	.09	.88*	.13	.00	.00
2	.22	.26	.35*	.17	.04	.04	.88*	.04
3	.35	.22	.22*	.22	.04	.13	.75*	.08
4	.39	.39*	.17	.04	.08	.63*	.17	.13
5	.22	.39	.39*	.00	.54	.29	.17*	.00
6	.00	.14	.59*	.27	.04	.04	.43*	.48
7	.18	.18	.23*	.41	.13	.08	.50*	.29
8	.35	.17	.35	.13*	.13	.21	.17	.50*
9	.35	.17	.35	.13*	.13	.21	.17	.50*
10	.17*	.26	.13	.43	.29*	.08	.08	.54
11	.04	.83*	.04	.09	.04	.92*	.00	.04
12	.22	.48*	.13	.17	.04	.65*	.13	.17
13	.52	.17	.17	.13*	.08	.04	.25	.63*
14	.39*	.26	.22	.13	.50*	.08	.08	.33
15	.13	.13	.52	.22*	.23	.13	.27	.37*
16	.09	.35	.48*	.09	.23	.23	.30*	.23
17	.43	.39*	.00	.17	.17	.63*	.17	.03
18	.26	.70*	.30	.00	.07	.73*	.17	.03
19	.30	.22	.17	.30*	.13	.13	.13	.60*
20	.09	.52	.35*	.04	.00	.10	.77*	.13
21	.00	.70*	.30	.00	.07	.73*	.17	.03
22	.30	.09	.17	.43*	.20	.17	.13	.50*
23	.43*	.04	.13	.39	.47*	.07	.47	.00
24	.09	.87*	.00	.04	.03	.97*	.00	.00
25	.22	.17	.13	.48*	.21	.31	.38	.10*
26	.22	.09	.09	.61*	.07	.17	.03	.72*
27	.09	.17	.13	.61*	.28	.14	.14	.45*
28	.13	.22	.48*	.17	.28	.10	.45*	.17
29	.26*	.43	.22	.09	.45*	.24	.21	.10
30	.09*	.39	.17	.35	.52*	.07	.14	.28
31	.00	.22	.35	.43*	.07	.10	.24	.59*
32	.17	.09	.13	.61*	.03	.14	.10	.72*
33	.73*	.23	.00	.05	.79*	.10	.10	.00
34	.36	.18	.27	.18*	.17	.17	.03	.62*
35	.09	.05	.23	.64*	.17	.07	.07	.69*
36	.18	.23*	.45	.14	.25	.43*	.21	.11
37	.00	.55*	.36	.09	.04	.50*	.36	.11
38	.18	.36*	.09	.36	.11	.64*	.18	.07
39	.50	.23	.09	.18*	.19	.19	.11	.52*
40	.14*	.45	.14	.27	.19*	.37	.17	.26
41	.23	.14*	.55	.09	.04	.46*	.08	.42
42	.23*	.36	.05	.36	.44*	.11	.22	.22
43	.27	.27	.09	.36*	.41	.22	.11	.26*
44	.18*	.32	.32	.18	.31*	.31	.19	.19
45	.36	.27	.14	.23*	.23	.18	.32	.27*

*Indicates correct answer

TABLE 17: Proportion of All Responses to Each Response Option on the Level Three Mathematics Computation Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.87*	.04	.09	.00	.93*	.00	.00	.07
2	.26	.09	.00	.65*	.10	.27	.03	.60*
3	.41*	.23	.32	.05	.30*	.17	.40	.13
4	.09	.41	.18*	.32	.10	.23	.30*	.37
5	.17	.39	.43*	.00	.10	.27	.40*	.23
6	.26	.04	.13	.57*	.47	.03	.10	.40*
7	.17	.78*	.04	.00	.20	.60*	.17	.03
8	.13	.87*	.00	.00	.20	.70*	.07	.03
9	.09	.39	.43*	.09	.20	.13	.47*	.20
10	.26	.57*	.09	.09	.27	.47*	.20	.07
11	.17	.09	.22	.52*	.17	.27	.13	.43*
12	.70*	.09	.13	.09	.67*	.17	.10	.07
13	.04	.91*	.04	.00	.07	.87*	.03	.03
14	.17	.30*	.30	.22	.10	.40*	.27	.23
15	.43*	.30	.22	.04	.07	.07	.63*	.23
16	.00	.04	.91*	.04	.07	.07	.63*	.23
17	.26	.43	.17*	.13	.17	.53	.30*	.00
18	.13	.26*	.35	.26	.10	.23*	.37	.30
19	.09	.17*	.74	.00	.03	.13*	.80	.03
20	.05	.09	.73*	.14	.07	.10	.73*	.10
21	.09	.36	.18	.36*	.21	.17	.10	.52*
22	.00	.83*	.13	.04	.00	.73*	.10	.17
23	.04	.00	.87*	.09	.03	.07	.80*	.10
24	.61*	.04	.30	.04	.57*	.13	.20	.10
25	.04	.09	.04	.83*	.03	.20	.03	.73*
26	.05	.18	.09	.68*	.03	.17	.10	.69*
27	.09	.83*	.00	.09	.10	.70*	.10	.10
28	.91*	.09	.00	.00	.83*	.10	.07	.00
29	.52*	.22	.22	.04	.47*	.30	.23	.00
30	.04	.30	.61	.04	.07	.30*	.43	.20
31	.23*	.00	.27	.50	.27*	.03	.40	.30
32	.26*	.13	.52	.09	.30*	.27	.40	.03
33	.05	.09	.73*	.14	.07	.07	.73*	.13
34	.05	.19	.14	.62*	.07	.10	.13	.70*
35	.21	.63*	.11	.05	.07	.60*	.30	.03
36	.11	.00	.16	.74*	.10	.07	.03	.80*
37	.05	.90*	.00	.05	.10	.67*	.03	.20
38	.11	.11	.42*	.37	.03	.14	.31*	.52
39	.00	.17	.28	.56*	.07	.18	.32	.43*
40	.17	.28	.00	.56*	.17	.17	.17	.50*
41	.11	.32	.11	.47*	.10	.60	.07	.23*
42	.12	.00	.76*	.12	.10	.10	.77*	.03
43	.18	.29	.18	.35*	.20	.40	.07	.33*
44	.41*	.29	.06	.24	.36*	.14	.18	.32
45	.12	.76*	.12	.00	.04	.79*	.07	.11
46	.59*	.12	.24	.06	.25*	.18	.50	.07
47	.24	.35	.18	.24	.39*	.25	.14	.21
48	.00	.69*	.06	.25	.07	.70*	.15	.07

*Indicates correct answer

TABLE 18: Proportion of All Responses to Each Response Option on the Level Three Mathematics Concepts and Applications Subtest

Item	English				Spanish			
	1	2	3	4	1	2	3	4
1	.05	.95*	.00	.00	.03	.73*	.10	.13
2	.05	.27*	.32	.36	.20	.23	.33	.23
3	.00	.09	.27	.64*	.03	.17	.07	.73*
4	.00	.41	.09	.50*	.03	.20	.37	.40*
5	.45*	.18	.27	.09	.23*	.27	.40	.10
6	.38	.14	.43*	.05	.40	.27	.27*	.07
7	.40	.30	.25*	.05	.40	.27	.27*	.07
8	.71*	.00	.14	.14	.50*	.07	.37	.07
9	.27*	.23	.09	.41	.43*	.17	.27	.13
10	.45*	.18	.09	.27	.43*	.17	.27	.13
11	.36*	.27	.27	.09	.40*	.23	.23	.13
12	.14	.05	.71*	.10	.10	.17	.57*	.17
13	.67*	.05	.14	.14	.47*	.10	.23	.20
14	.30*	.15	.30	.25	.27*	.17	.37	.20
15	.14	.38	.33*	.14	.17	.31	.41*	.10
16	.19	.14	.43*	.24	.23	.10	.47*	.20
17	.67*	.10	.19	.05	.57*	.30	.07	.07
18	.10	.33*	.52	.05	.07	.40*	.40	.13
19	.05	.33	.14	.48*	.33	.20	.13	.33
20	.14	.33	.38*	.14	.07	.17	.70*	.07
21	.00	.52*	.19	.29	.21	.45*	.14	.21
22	.14	.38	.43	.05*	.17	.23	.10	.50*
23	.14	.73*	.05	.09	.20	.73*	.02	.00
24	.14	.24	.33*	.29	.24	.31	.28*	.17
25	.43	.10	.33*	.14	.40	.13	.33*	.13
26	.05	.14	.19	.62*	.13	.10	.20	.57*
27	.82*	.00	.14	.05	.63*	.03	.27	.07
28	.10	.29	.62*	.00	.27	.17	.50*	.07
29	.05	.90*	.05	.00	.07*	.60	.27	.07
30	.10	.52*	.24	.14	.23*	.23	.27	.27
31	.33*	.24	.33	.10	.25	.14	.43	.18
32	.90*	.10	.00	.00	.77*	.13	.03	.07
33	.35	.30*	.25	.10	.45	.34*	.10	.10
34	.48	.14	.29*	.10	.59	.10	.21*	.10
35	.10	.29	.10*	.52	.07	.40	.10	.43
36	.25	.10	.15	.50*	.20	.07	.10	.63*
37	.62*	.05	.24	.10	.50*	.10	.10	.30
38	.05	.32*	.16	.47	.10	.47*	.10	.33
39	.05	.79	.05	.11*	.00	.43	.07	.50*
40	.11	.32*	.26	.32	.03	.57*	.23	.17
41	.35	.20	.15	.30*	.23	.17	.20	.40*
42	.16	.11	.58*	.16	.04	.11	.64*	.21
43	.42	.32	.21*	.05	.59	.28	.14*	.00
44	.11	.68*	.21	.00	.34	.45*	.14	.07
45	.06	.83*	.06	.06	.00	.70*	.13	.17
46	.06	.35	.18	.41*	.10	.14	.14	.62*
47	.11	.22	.22*	.44	.10	.14	.38*	.38
48	.00	.41	.41*	.18	.10	.43	.37*	.10
49	.25	.56*	.06	.13	.17	.41*	.17	.24
50	.11	.39*	.33	.17	.33	.30*	.27	.10

*Indicates correct answer

students found it easy to match, for example, casa to residencia and house to residence.

This particular example also points up another feature of several of the easiest items, namely, the presence of cognates. Examples of such pairs in the vocabulary subtest (item difficulties in parentheses) are:

1.	equivalent/equal	(.78)
	equivalente/iguales	(.75)
2.	affirmative/positive	(.67)
	afirmativo/positivo	(.75)

Such items were not always easy. Sometimes they fell in a mid-range of difficulty. Whenever this happened, the level of difficulty was about the same for each language. Examples are:

1.	dialogue/conversation	(.41)
	diálogo/conversación	(.54)
2.	confirmed/verified	(.43)
	confirmada/verificada	(.54)

In most cases of cognate pairs, the Spanish language item was somewhat easier than the English language item.

The items discussed above may be referred to as "pure" cognate pairs since both stimuli and responses are cognates. There were also examples of "partial" cognate pairs where only the stimuli were cognates. Examples of this type are:

1.	conclusion/end	(.43)
	conclusión/fin	(.46)
2.	terror/fear	(.52)
	terror/espanto	(.79)

Partial cognate pairs tended to be more or less difficult as a function of the frequency of the English response word. Thus while the first pair above was about equally difficult (.43/.46) the second pair was not (.52/.79). Occasionally, the difficulty levels of such item pairs was greatly discrepant as in the case of:

negotiated/bargained	(.21)
negociaron/comerciaron	(.87)

In those item pairs where cognates were present, there was a tendency for English and Spanish item scores to be approximately equal. Since the total test scores for the vocabulary subtest were not equal, we were interested in conducting an error analysis on those items where the discrepancy between English and Spanish item scores was larger. The

approach was first to identify the relevant English language items, then to examine the frequency of response to each of the distractors.

Two major patterns of wrong response occurred. In the case that a student did not know the right answer and there were no clues to such an answer, responses tended to be random, giving a typical response proportion of close to .25 for each option. There were cases, however, where the students went for one of the distractors in a strong way. In most of these cases, the option chosen bore a strong physical resemblance to the underlined stimulus word. Examples of such pairs are (proportion of response in parentheses):

1. conclusion/comparison (.35)
2. adequate/adjusted (.43)
3. data/dates (.35)
4. incident/accident (.26)
5. environment/investment (.39)
6. installment/implement (.38)
7. legitimate/legible (.37)
8. negotiated/negated (.32)

That students were paying attention to physical similarity between stimulus words and distractors is attested by the fact that random patterns of response occurred only on items where no distractor bore a physical resemblance to the stimulus word.

It may be useful here to comment on response tendencies on those items where the creation of a test by translation resulted in some awkward translations. Such items sometimes caused problems for those subjects who took the CTBS Espanol. At other times, they caused no trouble. Examples of the latter are:

1. personalidad magnética atractiva (.58)
2. charla tranquilizadora calmante (1.00)
3. reacción entusiasta ansiosa (.63)

Examples of the former are:

1. discurso estimulante incitante (.21)
2. descripción complicada enredada (.00)

There were other difficult Spanish items where the cause of the difficulty may have been more related to experience than to language per se. For example, in one item, "jolly old man" is given as the stimulus and "merry" the correct answer. The conjunction of these particular words immediately elicits the image of Santa Claus with whom the words "jolly" and "merry" are intimately associated. In the Spanish version, the stimulus reads "un anciano jovial" and the correct answer "alegre." Though Santa Claus (San Nicolas) is a familiar figure in Latin America, the term "anciano jovial" is not closely linked with him. Additionally, the word "jovial" is a low frequency word. The proportion of correct responses to the Spanish version was .26, to the English version .22 indicating that the U.S. Christmas associations are not known to our sample in English nor are they mirrored in Spanish. By contrast, the English language norm sample demonstrated a proportion of correct response of .77 on this item in English.

Other errors may be directly related to culture differences. For example, one item pair is given as:

man of good character/qualities (.29)
un hombre de buena reputación/fama (.21)

One of the response options on this item was "education/educación." The proportion of responses to "education" was .62 and to "educacion" was .50. A Latin American will quickly know why the students responded in this way. For them, there is an intimate link among character, reputation, and educación. They responded correctly from the standpoint of their culture even though they got the item wrong.

We may conclude the discussion of response tendencies on the vocabulary subtest by presenting a list of guessing strategies that the LEP students seem to have used on the CTBS/S when the correct answer was not surely known. These strategies are, of course, inferences. Some of them are based on several items. Some of them are based on fewer items. All of them are presented as hypotheses.

H_1: If the answer is not known but one distractor is within the semantic region, even if it is an antonym, choose it.

H_2: If the answer is not known but there is one distractor that is frequently used as a modifier of the stimulus word, choose it.

H_3: If the answer is not known but there is one distractor that experientially occurs in conjunction with the stimulus (e.g., accident, damage), choose it.

H_4: If the answer is not known but there is a distractor whose orthographic form is similar to that of the stimulus word, choose it.

H_5: If the answer is not known but one of the distractors is a cognate (or false cognate) of the stimulus word, choose it.

H_6: If the answer is not known but one of the distractors is known to mean the same as the whole item stem, choose it.

H_7: If the answer is not known and none of the "clues" listed in the first six hypotheses can be found, guess blind.

We can now move on to a consideration of findings on the Reading Comprehension subtest. The first step in the analysis was to look for an effect due to language of the test. A visual inspection of Tables 12 and 16, reveals that the English language version of the items was typically more difficult than the Spanish language version. This difference was examined statistically and was found to be significant. The MANOVA is presented later.

In order to learn as much more as possible, the item difficulties were organized by reading selection. Visual inspection of the arrays revealed that there were systematic differences in difficulty as a function of reading selection. Interestingly, the two most difficult selections on Level 1 were the two dealing with scientific subjects, one with the invention of a phonograph and the other with satellites. It is interesting to speculate that the difficulty may be attributed to lack of regular contact with the world of science in the home countries. At the same time, it must be admitted that we are unable to separate such a hypothesized cause from other possible causes rooted in language itself. The nature of the passages in Level 3 made it impossible to check the hypothesis. It was found, however, that in Level 3 also, there were clear differences in difficulty as a function of reading selection.

This analysis led to an even more interesting observation. This was the nature of the distribution of item difficulties within reading selections. This led to speculation that there might be systematic variability in item difficulty as a function of type of task. To check this out, each item was categorized in terms of the following task descriptions.

1. locating details
2. getting the main idea
3. establishing order of events
4. drawing inferences

5. determining meanings of words and expressions
6. summarizing sub-sections
7. selecting descriptive words

While a visual inspection of item difficulties tended to suggest that items requiring location of details were easiest, there was such a confounding of task difficulty with language difficulty that it was decided to try to reduce some of the "noise." Since it had been noted that the LEP students responding to the Spanish version earned lower item scores on some items than did FEP students responding to the English version, it seemed plausible to assume that on these particular items, the LEP students were experiencing some difficulty over and above any possible language difficulty. It was therefore decided to look at these items as a special subset. An item was selected for further investigation if the Spanish version/LEP student difficulty index was at least .16 lower than the English version/FEP student index. The selection of the criterion .16 was arbitrary.

At Level 1, there were 13 items out of 41 (items 42-45 were omitted from the analysis since the Spanish and English versions were based on different poems) that met the criterion. The distribution of these items by task type was as follows:

1. locating details (items 13, 19, 36, 39)
2. getting the main idea (items 15, 20)
3. establishing order of events (item 8)
4. drawing inferences (items 12, 27)
5. determining meaning of words and expressions (items 18, 25, 32)
6. summarizing sub-sections (item 14)
7. selecting descriptive words (item 22)

Note that "locating details" accounts for 30 percent of the items in this subset while items requiring this skill make up 45 percent of all items in the subtest (items 42-45 omitted). This may indicate that there are indeed "task" factors that create the observed increased difficulty.

The reader should not infer from the above that the study sample never had difficulty in locating details nor that they always had difficulty making inferences, etc. The finding is merely that, on those items which LEP students found excessively more difficult than did FEP students, the tasks were most likely to require more cognitively demanding skills. These findings may be interpretable in terms of Cummins' recent formulation where the difficulty of a given language task is a function of the interaction of two factors,

namely, its contextual embeddedness and its cognitive demandingness (Cummins, 1981). We shall discuss this hypothesis in greater detail in the discussion section of the report.

It would have been preferable if the generalizations about item difficulties as a function of the interaction of linguistic and task factors could have been checked on Level 3 of the test. This was not possible since item difficulties on that test were based on a mixed sixth and eighth grade study sample. This altered item difficulties in the study sample such that comparisons with the norming sample were questionable.

We turn now to an analysis of items on the Mathematics Concepts and Applications subtest. (An analysis of items on the Mathematics Computation subtest was not done since item difficulties on the Spanish and English versions were essentially equal. This could be attributed to the fact that language proficiency plays no significant role in the latter subtest).

In contrast, language proficiency would intuitively seem to play an important role in responding to the Mathematics Concepts and Applications subtest. As we shall see later, differences in subtest scores as a function of the language of the test were significant at Level 1 only though even on Level 3, the absolute value of the mean on the English version was lower than on the Spanish version.

The analysis of item difficulties on the Concepts and Applications subtest was conducted in a manner similar to the analysis of item difficulties on the Reading Comprehension subtest. The appropriateness of this analysis was suggested by the fact that on many items, LEP students taking the Spanish version did as well as FEP students taking the English version. On other items, the LEP students did substantially more poorly. This suggested, once again, the study sample might be experiencing difficulty over and above that due to language alone.

Using the .16 criterion, a subset of 23 Level 1 items was selected for special examination. The following list gives the kinds of conceptual areas on which LEP students did more poorly than FEP students regardless of the language of the test (unless otherwise indicated):

1. place value
2. reading time from a clock face
3. determining the number of U.S. coins required to make a given sum
4. working with the distributive and commutative properties of numbers
5. distinguishing between odd and even numbers (problem on English version only)

6. comparing numbers of items in two sets
7. combining two geometric figures to create a third figure
8. determining calendar time
9. determining lengths of objects in feet and inches (most difficult in English)
10. determining unknowns in an equality
11. determining the missing number in a number series
12. computing values in word problems

The finding for (12) was surprising. One might have guessed that there would be a strong effect due to language in word problems. Such an effect did not appear. LEP students did equally poorly on these items regardless of the language of the test (only one exception to this in eight items).

A similar analysis for Level 3 was not conducted because of the problem of combining data for sixth and eighth graders.

An Analysis of Effects

Separate but parallel analyses were conducted for Level 1 and Level 3 of the CTBS. These analyses included an analysis of scale reliabilities as well as a multivariate and univariate analysis of variance. Two multivariate analyses were performed. First, "reading" was investigated using reading vocabulary and reading comprehension as dependent variables, and then "mathematics" was examined with mathematics computation and mathematics concepts and applications as dependent variables. The results of these analyses are presented in this section, while their interpretation is contained in the subsequent section.

Reliability Analysis

One of the important concerns for this study was the psychometric properties of the English and Spanish versions of the CTBS. These properties were investigated both empirically and judgmentally, and at the item and scale levels. The results of the judgmental and item analyses have been presented in a preceding section of the report. This section discusses the result of the analysis of scale properties.

Each of the four test scales used in this study was submitted to an analysis of its psychometric properties. The analysis was performed using the SPSS software routine entitled "Reliability." This procedure provides a thorough item and scale analysis employing a variety of test score models. In the present case, we used a traditional test

theory model with coefficient alpha as the measure of scale reliability. This coefficient is a measure of internal consistency for the items of the scale and represents the average split-half reliability for all possible combinations of items. The results from these analyses are reported in Tables 19-20. As mentioned previously, separate analyses were conducted for the English and Spanish versions of the CTBS.

An examination of the Level 1 results has two important features. First, the reliabilities for all four scales are acceptably high (ranging from .82 to .96) and are very comparable for the two versions of the test. This would seem to indicate that the scales are fairly homogeneous at this level and across forms. Second, the performance on the English version of the test is considerably lower than the Spanish version. This discrepancy is largest for the subtests involving language (Vocabulary, Reading Comprehension, and Mathematics Concepts and Applications) and smallest in the pure non-verbal subtests, that is, Mathematics Computation. Since comparable groups of students received the two forms of the test, this finding indicates a clear difficulty difference in the two forms.

The Level 3 results are substantially the same as those of Level 1. Again, the English version is considerably more difficult for the students. However, the reliabilities for the Vocabulary and Reading Comprehension subtests on the English form are considerably lower than those on the Spanish form. The Reading Comprehension reliability of .55 is very low. This may mean that this subtest is not factorially pure and that more than a single ability is being measured here. Further, this lack of reliability will probably make it more difficult to find treatment effects at this level.

Overall, these results are striking in the difference in the difficulty levels of the two test versions. This is particularly true since the tests are direct translations and the students are of comparable ability.

Analysis of Variance Results: Level 1

The results of the analysis for Level 1 of the CTBS are presented in Table 21 for reading and Table 22 for mathematics. Examination of these results shows the presence of several statistically significant sources of variance. In the case of reading, and its component subtests, there was a very strong effect ($p < .001$) present for the test language. In addition, the method of test administration also was statistically significant ($p < .01$) both univariately and multivariately. Finally, the interaction of test language and test administration was also statistically significant ($p < .05$). Examination of the cell means contained in Table

TABLE 19: Scale Reliabilities, Averages and Dispersions for All Students*

		Level 1		Level 3	
		English	Spanish	English	Spanish
Vocabulary (40 items)	$\bar{X}$	14.23	23.15	14.42	23.80
	SD	6.84	6.21	5.70	8.00
	α	.84	.82	.75	.89
Reading Comprehension (45 items)	$\bar{X}$	17.15	24.28	16.84	23.40
	SD	8.61	9.81	4.50	8.54
	α	.88	.91	.55	.89
Math Computations (48 items)	$\bar{X}$	34.31	37.75	22.87	25.30
	SD	9.84	11.45	10.68	9.99
	α	.93	.96	.93	.92
Math Concepts (50 items)	$\bar{X}$	22.50	29.50	18.42	22.40
	SD	8.84	10.82	10.13	9.68
	α	.88	.93	.92	.90

*Treats omits as incorrect

TABLE 20: Scale Reliabilities, Averages and Dispersion for Students Who Completed Each Scale Item

		Level 1		Level 3	
		English	Spanish	English	Spanish
Vocabulary (40 items)	$\bar{X}$	16.59	24.23	15.27	24.50
	SD	6.57	5.22	5.69	8.11
	α	.82	.75	.74	.90
Reading Comprehension (45 items)	$\bar{X}$	21.37	31.56	18.96	27.14
	SD	8.08	6.99	4.53	8.84
	α	.84	.81	.50	.89
Math Computations (48 items)	$\bar{X}$	37.32	41.00	25.90	25.56
	SD	7.25	5.23	8.92	9.34
	α	.88	.83	.89	.90
Math Concepts (50 items	$\bar{X}$	27.67	33.12	26.00	25.88
	SD	6.97	8.15	7.90	9.71
	α	.80	.87	.82	.87

TABLE 21: ANOVA for Level One CTBS
Reading Vocabulary Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	10.83	0.31	0.5820
Language	1	1324.35	37.35	0.0001
TW* Language	1	159.18	4.49	0.0371
Administration	1	201.18	5.67	0.0195
TW* Administration	0	0.00	----	------
Language* Administration	1	287.72	8.11	0.0056
TW* Language* Administration	0	0.00	----	------

ANOVA for Level One CTBS Reading Comprehension Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	223.53	3.47	0.0600
Language	1	874.15	13.58	0.0004
TW* Language	1	135.25	2.10	0.1510
Administration	1	1128.72	17.54	0.0001
TW* Administration	0	0.00	----	------
Language* Administration	1	335.38	5.21	0.0250
TW* Language* Administration	0	0.00	----	------

TABLE 22: ANOVA for Level One CTBS
Mathematics Computation Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	548.12	6.20	0.0148
Language	1	170.23	1.92	0.1691
TW* Language	1	37.45	0.42	0.5171
Administration	1	237.79	2.69	0.1040
TW* Administration	0	0.00	----	------
Language* Administration	1	412.33	4.66	0.0338
TW* Language* Administration	0	0.00	----	------

ANOVA for Level One CTBS Mathematics Concepts and Applications Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	40.79	0.49	0.4878
Language	1	679.77	8.10	0.0056
TW* Language	1	5.05	0.06	0.8060
Administration	1	152.18	1.81	0.1819
TW* Administration	0	0.00	----	------
Language* Administration	1	35.09	0.42	0.5197
TW* Language* Administration	0	0.00	----	------

24, clearly shows higher levels of performance for students who received the Spanish version of the CTBS and who were administered the test orally. No statistically significant effects associated with the testwiseness training were identified here, though the multivariate and univariate results for reading comprehension were in a positive direction.

The results for mathematics were similar but not quite as dramatic as those for reading. Again, test language and the language by administration interaction term were statistically significant at the multivariate level. However, examination of the univariate results shows the language effects to be confined to the Mathematics Concepts and Application subtest, while the interaction term is significant only for the Mathematics Computation subtest. Of interest here, is the presence of a statistically significant ($p < .05$) effect associated with testwiseness training. This result obtained both at the multivariate level and at the univariate level. Examination of the cell means in Tables 23-24 shows that students who received testwiseness training tended to perform somewhat higher on the Mathematics Computation subtest.

Overall, these results suggest the presence of a strong test language and test administration factor, particularly in the verbal areas. Additionally, there was some indication that testwiseness was influencing performance, though its strength was considerably less than that of the other factors.

Analysis of Variance Results: Level 3

The results of the Level 3 analysis are presented in Table 25 for the Reading subtest and in Table 26 for the mathematics subtest. Examination of these tables shows that the Level 3 results were statistically non-significant for the most part. The only significant effect was that of test language for the reading (multivariate) and its component subtests. No other main effects or interactions were statistically significant. The results of the mathematics subtests produced no statistically significant results. Overall, then, the Level 3 results were confined to a statistically significant test language effect for the reading subtests. These results correspond to that found at Level 1, though other factors were also statistically significant at that lower level.

A general note of caution concerning these results for both test levels needs to be made. Specifically, the relatively small sample size and the low test performance (often approaching the chance level) probably makes these findings unstable. Thus they should be viewed as suggestive rather than definitive.

TABLE 23: Means for Level One CTBS
Tested by Main Effects ANOVA

Variable	Condition	Reading Vocabulary	Reading Comp.	Math Comp.	Math Concepts
TW Training	No (N=22)	17.82	20.64	40.59	26.05
	Yes (N=66)	19.64	21.47	36.39	26.53
Language of Test	Spanish (N=51)	22.67	24.06	38.55	28.90
	English (N=37)	14.38	17.41	35.92	22.97
Admin. Mode	Oral (N=33)	20.97	25.21	38.88	28.24
	Standard (N=55)	18.11	18.89	36.58	25.31

TABLE 24: Means for Level One CTBS Tested by Three-Way ANOVA

Cell			Reading Voc.	Reading Comp.	Math Comp.	Math Concepts
TW Training	Language of Test	Admin. Mode				
TW	Spanish	Oral (N-20)	22.80	26.20	42.25	31.45
		Standard (N=20)	23.50	22.35	33.25	26.85
	English	Oral (N=13)	18.15	23.69	33.69	23.31
		Standard (N=13)	10.31	10.62	34.92	21.69
No TW	Spanish	Oral (N=0)	----	----	----	----
		Standard (N=11)	20.91	23.27	41.45	28.00
	English	Oral (N=0)	----	----	----	----
		Standard (N=11)	14.73	18.00	39.73	24.09

TABLE 25: ANOVA for Level Three CTBS
Reading Vocabulary Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	0.12	0.00	0.9618
Language	1	1561.68	29.04	0.0001
TW* Language	1	50.63	0.94	0.3355
Administration	1	6.56	0.12	0.7280
TW* Administration	1	.3.36	0.06	0.8034
Language* Administration	1	61.18	1.14	0.2901
TW* Language* Administration	1	86.10	1.60	0.2103

ANOVA for Level Three CTBS Reading Comprehension Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	3.46	0.06	0.7998
Language	1	586.62	11.00	0.0015
TW* Language	1	0.05	0.00	0.9760
Administration	1	104.84	1.97	0.1656
TW* Administration	1	10.70	0.20	0.6556
Language* Administration	1	3.86	0.07	0.7886
TW* Language* Administration	1	79.54	1.49	0.2284

TABLE 26: ANOVA for Level Three CTBS Mathematics Computation Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	0.15	0.00	0.9718
Language	1	11.82	0.10	0.7532
TW* Language	1	0.32	0.00	0.9585
Administration	1	0.79	0.01	0.9353
TW* Administration	1	20.95	0.18	0.6757
Language* Administration	1	8.63	0.07	0.7882
TW* Language* Administration	1	171.43	1.45	0.2337

ANOVA for Level Three CTBS Mathematics Concepts and Applications Subtest

Source	DF	SS	F	Pr > F
Testwiseness	1	7.87	0.07	0.7896
Language	1	178.15	1.63	0.2086
TW* Language	1	72.20	0.66	0.4199
Administration	1	146.29	1.34	0.2521
TW* Administration	1	0.84	0.01	0.9306
Language* Administration	1	6.18	0.06	0.8130
TW* Language* Administration	1	258.56	2.36	0.1293

TABLE 27: Means for Level Three CTBS Tested by Main Effects ANOVA

Variable	Condition	Reading Vocabulary	Reading Comp.	Math Comp.	Math Concepts
Testwiseness Training	No (N=32)	19.25	19.84	23.59	19.38
	Yes (N=41)	18.58	19.51	23.00	18.90
Language of Test	Spanish (N=36)	23.67	22.61	23.31	20.33
	English (N=37)	14.22	16.78	23.22	17.92
Admin. Mode	Oral (N=32)	19.75	21.28	23.38	20.97
	Standard (N=41)	18.20	18.39	23.17	17.66

TABLE 28: Means for Level Three CTBS
Tested by Three-Way ANOVA

Cell			Reading Voc.	Reading Comp.	Math Comp.	Math Concepts
TW Training	Language of Test	Admin. Mode				
TW	Spanish	Oral (N=11)	21.45	22.27	22.64	19.27
		Standard (N=8)	24.63	23.25	24.25	19.88
	English	Oral (N=11)	17.64	19.00	24.45	21.45
		Standard (N=11)	12.27	14.55	21.00	15.27
No TW	Spanish	Oral (N=6)	25.00	25.83	25.00	26.33
		Standard (N=11)	24.45	20.73	22.36	18.45
	English	Oral (N=4)	13.00	18.00	20.00	16.25
		Standard (N=11)	13.18	16.36	25.36	17.64

DISCUSSION AND CONCLUSIONS

Effects of Culture

It is often claimed that norm-referenced English language tests like, for example, the CTBS are culturally biased. This hypothesis is sometimes advanced to explain the relatively lower performance of language minority students on such tests.

While the claim may be true, it is seldom accompanied by a systematic definition of the key term "culture." Apart from such a definition, the claim has little explanatory value. That is, there is nothing in the claim itself to help one to identify and correct specific offending cultural features. Since the aim of this study is the improvement of estimates of achievement in students with limited English proficiency, it was incumbent on us to provide a definition of the term "culture" that would allow us to (1) evaluate the validity of the claim and (2) to recommend some action to correct it. We turn now to the presentation of such a definition.

According to Goodenough, "culture consists of the various standards for perceiving, believing, valuing, and doing... that [a person] attributes to other persons as a result of their actions and admonitions. By our definition of culture, the standards that a person thus attributes to a particular set of others are for him the culture of that set. (Cited in Gibson, 1976, p. 9)."

With this definition, Goodenough places the locus of culture "inside" the person. It is important to recognize this point since most popular conceptions place the locus of culture "outside" the person, usually in the form of modes of dress, diet, customs, language, etc. The naivete of this latter view of culture can be perceived by considering the following example. Suppose an Anglo-American should decide to begin to dress like a Mexican (but what does a representative Mexican wear?), eat what Mexicans eat (and what does a representative Mexican eat?), behave like a Mexican (and how does a representative Mexican behave?), speak Spanish (and what about Mexicans who speak Indian dialects, German, Chinese, etc.), etc. Would our Anglo-American have become culturally a Mexican? Clearly he would not. Culturally, he would be exactly what he was before but now he would be masquerading in terms of his Mexican stereotypes, none of which have anything at all to do with what makes a Mexican what he culturally is. In fact, by our definition of culture, the term "Mexican" has little to do with culture. Rather, the term simply describes a person's nationality.

For our purposes, the value of this definition of culture lies in the assistance it gives in identifying examples of culture bias as well as in making distinctions between culture bias per se and other non-cultural features of a test which nonetheless systematically affect a student's ability to make correct responses. By the definition we have adopted, we shall say that an item is culturally biased if getting the correct answer requires that the examinee be competent in a culture in which he/she is not competent or, more precisely, be able to function in terms of the standards for perceiving, believing, valuing, and doing that define the culture.

Using this definition as our standard, we looked for and found little of culture bias in the CTBS. In the entire item analysis of Levels One and Three, only one item was found that might be said to be culturally biased. All other cases in which the study subjects showed a strong tendency to choose an answer other than the correct one, factors other than culture seem to have been involved. We therefore find the manufacturer justified in the claim that CTBS scores are affected minimally by culture bias.

Effects of Language

Clearly (and not surprisingly), the language of the test was a more powerful determiner of test scores than was culture. The students did much better on the Spanish version than they did on the English version.

Parenthetically, the reader should note that we treat language alongside and independent of culture, not as an element of culture itself. This is done because the definition of culture that has been adopted does not permit its inclusion. The correctness of this decision can be appreciated by considering the fact that groups having widely divergent cultures often speak the same language. While language may be used as a medium for communicating or revealing culture, it cannot be identified as a part of culture itself.

Returning now to the discussion of the findings, it was found that the subtests can be ranked by difficulty in terms of their dependence on academic English proficiency. The Reading Vocabulary and Reading Comprehension subtests were the most linguistically demanding and therefore the most difficult, the Mathematics Concepts and Applications subtest was the next most linguistically demanding and therefore the next most difficult, and the Mathematics Computation subtest was the least linguistically demanding and therefore the least difficult. In fact, with regard to the latter subtest, we could distinguish no difference in the level of difficulty between the English and Spanish versions. We conclude,

therefore, that for the non-English-proficient student, the more a subtest depends on English language proficiency, the more difficult it becomes.

As we pointed out earlier, this finding only states the obvious. Its importance lies perhaps in (1) documenting the magnitude of an effect whose presence is known to every bilingual educator and (2) in leading us to ask further questions about the specific nature of the language difficulties that language minority students encounter when they take tests like the CTBS in English.

We were led in our analysis by theoretical work done by Cummins. In his earlier theorizing, Cummins' research led him to distinguish between conversational skills (Basic Interpersonal Communication Skills, i.e., BICS) and and academic language skills (Cognitive-Academic Language Proficiency, i.e., CALP). He posited distinct growth curves for each of these two types of language proficiencies. BICS was assumed to develop rapidly from birth to age five, then to level off, reaching asymptote later in life. Growth in BICS was assumed to be a function of language acquisition (as opposed to language learning).

In contrast, CALP was assumed to develop less rapidly to about age fifteen, then to level off, reaching asymptote later in life. Growth in CALP was assumed to be a function of learning (as opposed to language acquisition) (Cummins, 1980).

The importance of the distinction between BICS and CALP for this study lies in Cummins' assumptions about the relation between BICS and CALP in learners of a second language. BICS in a second language is assumed to develop first. Students who have good BICS development are often quite proficient in the second language. This, however, is frequently misleading since full proficiency means that one is proficient not only in second language BICS but also in second language CALP. Students who lack proficiency in CALP, even though they are proficient in BICS, may yet fail tests like the CTBS, which are tests of CALP.

Since that first formulation, Cummins has carried his theorizing further (Cummins, 1981). In place of the two relatively discrete language functions (BICS and CALP), he posits a two dimensional surface on which the dimensions may be named "level of cognitive demand" and "level of contextual embeddedness." Level of cognitive demand is continuously distributed between tasks that are cognitively undemanding at one extreme to tasks that are cognitively demanding at the other. Level of contextual embeddedness is continuously distributed between tasks, that are context-embedded at one extreme to tasks that are context-reduced at the other. The crossing of these two continuums defines four quadrants. The quadrants define four classes of communication as follows:

1. Quadrant A: communications that are cognitively undemanding and context-embedded
2. Quadrant B: communications that are cognitively demanding and context-embedded
3. Quadrant C: communications that are cognitively undemanding and context-reduced
4. Quadrant D: communications that are cognitively demanding and context-reduced

Communications in Quadrant A are characterized by the fact that "participants can actively negotiate meaning... and the language is supported by a wide range of paralinguistic and situational cues" on the one hand and by the fact that "the lingistic tools have become largely automatized (mastered) and thus require little active cognitive involvement for appropriate performance on the other." These kinds of communications are relatively easy. They include much of Cummins formerly called BICS.

Communications in Quadrant D are much more difficult. They are characterized by heavy reliance "on linguistic cues to meaning" on the one hand and by the fact that the linguistic tools have not become automatized and thus require active cognitive involvement on the other. Characteristics of communications in Quadrants B and C may be inferred from what has been said about Quadrants A and D.

We have taken some time to describe Cummins' work because we found it useful first in suggesting an approach to item analysis and second in interpreting the results.

With respect to the CTBS, it seemed useful to think of all communications in the battery as falling in Quadrants C and D since the test format provides no opportunity for negotiation of meaning, no paralinguistic or situational cues and, consequently forces a total dependence on linguistic cues to meaning. For a student with limited English proficiency such tasks in English must necessarily be difficult.

Even so, Cummins' model would predict that such tasks would vary in difficulty as a function of the other dimension of the model, that is, as a function of cognitive demandingness. We believe that the item analysis data are consistent with the prediction (though certain non-linguistic features that are not a part of the model almost certainly contributed to variations in item difficulty).

We may conclude this part of the discussion by saying that in all probability, a major part of the difference in difficulty between the English and Spanish versions of the

CTBS for the students in this study lay in their lesser ability to handle in English the context reduction that is an integral part of the battery.

Effects of Providing Oral Assistance

Some of the students in this study were given an oral administration of the Reading Vocabulary, the Reading Comprehension, and the Mathematics Concepts and Applications subtests. An analysis of the effects of this treatment revealed no main effects. However, administration procedures did interact with test language, thus providing a significant advantage for students who received the oral administration procedure in combination with the English language version. This effect occurred on the Level One Reading Vocabulary and Reading Comprehension subtests.

This interaction is easily comprehensible since one would expect that the target population would be less proficient in reading English than in reading Spanish. The fact that it occurred with fourth grade students but not with sixth and eighth grade students may indicate that among sixth and eighth grade students, though their overall proficiency with context-reduced English was poor. their reading skills and aural skills were about equal for this kind of English while for the fourth grade students, their aural skills were ahead of their reading skills for this kind of English. Thus, having items read to them orally helped them to get some items right that they would have missed otherwise.

It should be noted in passing that, because there were no fourth grade students in the treatment conditions defined as "no testwiseness/Spanish version/ oral administration" and "no testwiseness/English version/oral administration," the language by administration interaction could be tested only for students who had received testwiseness training. If one so restricted the analysis of Level Three data in exactly the same way, the interaction may have been significant at Level Three also. The reader can satisfy himself/herself on this by checking that portion of Table 28 dealing with students who had had testwiseness training. The differences were clearly in favor of students who had received the English version and the oral administration. This seems to point to a nascent possibility that the oral administration works best for students receiving the English version if they have also had testwiseness training. The direction of the data favor such a hypothesis but at Level 3, the three-way interaction was not significant.

We conclude, on the basis of these considerations that those students whose English oral skills are ahead of their English reading skills may get more items correct if the

items are read to them. Such a test naturally ceases to be a test of reading and becomes a test of ability to comprehend orally-presented context-reduced English. This finding might have important implications for diagnostic procedures in bilingual special education. The sampling limitations in this study prevent us from being more confident on this issue.

Effects of Testwiseness Training

Testwiseness training did not have the power that was anticipated. Since the training was directly relevant to the tasks required by the test and since the examinees who took the training appear not to have benefitted from it. we must assume either that all the examinees already possessed the relevant skills or that none of them did. If the latter was the case (and it appears to have been), then the testwiseness training procedures and materials failed to accomplish their objectives. Assuming that this is true, it is probably correct to assume that the failure was due to insufficient amounts of practice on the skills referenced in the training materials. It would appear, therefore, that any re-work of those materials should incorporate more practice exercises. Of course, it would be even more effective if teachers would incorporate these types of tasks into mainstream instruction. Students would then already be well-prepared to do tasks of those types, thus obviating the need for some parts of the testwiseness training materials.

Other modifications in the testwiseness training materials were suggested by the item analyses. These analyses showed that the students exhibited certain systematic response tendencies that were counterproductive. It should be possible to take advantage of this information, using it to modify the testwiseness training procedure, including specific instructional sequences designed to eliminate these tendencies and to replace them with more productive ones. If this were done, there is a high probability that testwiseness training would result in significant differences in favor of students receiving such training.

Effects of Using a Translated Test

The CTBS Espanol is a translation of the CTBS, Form S in English. The translation was done by educators from the Norwalk-La Mirada (California) Unified School District. For the manufacturer, the translation was a fortuitous gift since the nature of the market would probably have discouraged the manufacturer from producing such a test on its own. For the

field of bilingual education, it was equally fortuitous since it has permitted bilingual program personnel to document learning that might otherwise have gone undocumented.

Nevertheless, the creation of a Spanish language achievement test by translation brought certain difficulties with it. In defense of the translators, it must be recognized that the items which might have created the most difficulty in translation were not translated at all. They were replaced by new items written in Spanish only. Thus, it is incorrect to say that the CTBS Espanol is totally a translation from English.

Having said that, certain problems remain. The most apparent problem might be called the problem of awkwardness. The translation from English to Spanish sometimes resulted in the use of Spanish expressions which are not castizo, that is, expressions that would hardly, if ever, appear in common written Spanish and, as a consequence, have an awkward feel to them.

A second problem has to do with word frequency. It was the impression of the socio-linguist who contributed to the item analysis that English words having a relatively high frequency were translated using Spanish words with a relatively low frequency. Since it was outside the scope of this study to check relative word frequencies, it remains as a hypothesis. Nevertheless, the principal investigator and the socio-linguist both feel strongly that a careful investigation would support the hypothesis. If the hypothesis were confirmed, there would be good grounds for claiming that the Spanish version is more difficult than the English version.

A third problem appeared in Level B of the test, a level that was not used in the conduct of this study. In the Word Recognition I subtest, distractors were created using both meaningful words and non-sense words. The Spanish version never used non-sense words. This is a non-parallel feature of the two versions which unwittingly introduced phonetic "traps" in the English version. For example in Item D1 of the Word Recognition II subtest, we find a picture of a foot. The students are supposed to mark the word which names that picture. The four options are fut, foot, look, and fish. The correct answer is of course the second option. The three distractors all look reasonable as long as we are thinking in terms of a student population who have English as the primary language. As soon, however, as we come to students with a primary language of Spanish we find that the first distractor may become almost irresistably attractive. This is simply because of the phonetic character of the Spanish language and the pronunciation of the word fut as in, for example, the word fútbol. Any Spanish speaking student who is unfamiliar with the intricacies of the English vowel system is likely to be irresistably attracted

to fut and therefore to get a wrong answer. A similar kind of problem is found in Item 4 of the same subtest. Here we have a picture of a hand. The response options are hand, han, have, bunt. For English speakers, the answer is clear. For Spanish speakers, it is not at all clear since in Spanish no word ever ends in nd. It is known from experience that Spanish speakers who are pronouncing English words of this type almost inevitably drop the final d. It is probably safe to assume that the reason for dropping the d is that they simply do not hear it. Given this situation, the Spanish speaker may be almost irresistably drawn to the option han which is an incorrect answer. Numerous other problems of a similar nature can be identified.

A fourth problem has to do with the possibility that two words that are dictionary translations of one another, say house and casa, are psychologically and semantically distinct. Existing studies of the associative structures of such words in balanced bilinguals have shown that this is often the case.

The import of these observations is that the CTBS, Form S and the CTBS Espanol may not be parallel at all. Their factor structures may be quite different. Even so, the CTBS Espanol represents a careful piece of work. In the absence of a non-translated Spanish language test of equal quality, the use of the CTBS Espanol can be highly beneficial.

Effects of Unfamiliar Item Content

Being born and raised in one place automatically means that one is likely to become familiar with the environment associated with that place. It also means that one is unlikely to become familiar with the important features of environments with which one has had no previous direct experience. If the features of the unknown environments are incorporated into the content of test items, they are likely to result in negative bias for those examinees for whom the item content is unfamiliar. This kind of difficulty seems to have occurred with a few items. For example, some students had difficulties with items that referenced U.S. coins. This is understandable for the population in question. Such questions clearly favor those examinees who have handled such coins on a regular basis.

It is impossible to tell from the data produced by this sutdy the number of items over which the examinees stumbled simply because they had had no prior experience with the things referenced in the test items. There may have been many such items.

Effects of Prior Training

There was a significant number of items that were equally hard in both English and Spanish. On these same items, English-speaking students in the norm-sample did much better. In these cases, given that neither language nor culture can be invoked as an explanation for their comparatively poor performance, one must assume that in their prior schooling, these students simply did not learn how to solve these kinds of problems. It is therefore incumbent on teachers of these students to take note of the classes of tasks which the students find hard in either language and to provide them with specific appropriate instruction and practice.

Limitations of the Study

The greatest limitation of the study lies in the sample size. It may be that in a larger sample study, some of the non-significant effects would have been significant.

A second limitation seems to have been the inadequate numbers of practice exercises in the testwiseness training materials that could lead the students to the application of those same skills in the test-taking situation. To correct this situation, it would be necessary to add more practice exercises and to expand accordingly the number of training sessions.

APPENDIX

"Teacher's Manual for Training
in Testwiseness"

TEACHER'S MANUAL FOR
TRAINING IN TESTWISENESS

A Training Program Adapted for Use
With the Comprehensive Tests of
Basic Skills (CTBS/S)

English Version

Developed by:
Applied Social Research
P.O. Box 158
Pacific Palisades, California
(213)454-2104

ACKNOWLEDGEMENTS

The development of this training program began with a search for relevant literature. This search was conducted by Beverley Greer. Items that she located were turned over to Drs. William Doherty and Nancy Russo. They used the literature to extract the features which previous research had indicated should be incorporated into a training program. Those features were used to develop a set of specifications. These specifications were given to Dr. Lois Weinberg who in turn wrote this Manual and developed all the related practice materials. The program was field tested with fourth, sixth, and eighth grade students in La Habra Elementary School District. The field test revealed a need for minor modifications. The modifications have been made. The present document is the result. It, together with the practice materials were translated into Spanish by Dr. Richard Piper.

TO THE TEACHER

The reason for this program

"Training in Testwiseness" was developed as part of a study entitled "Improving Estimates of Achievement in Students with Limited English Proficiency." The study was funded under a cost-sharing grant awarded to Applied Social Research by the National Institute of Education. It was dedicated to finding some ways to help limited English proficient (LEP) students to earn scores that would more accurately reflect their true levels of achievement. Anyone who would like to have a copy of the study should ask for it by name. Call or mail your request to:

Applied Social Research
15219 Sunset Boulevard, Suite 201
Pacific Palisades, CA 90272
(213) 454-8464 - (213) 459-4264

The goals and objectives of this program

Many LEP students come to this country not knowing how to take tests like the Comprehensive Tests of Basic Skills (CTBS). They have not learned the strategies that knowledgeable students use regularly to improve their scores. This places the LEP students at a disadvantage vis a vis these other students.

The goal of the "Training in Testwiseness" program is to erase this disadvantage, to teach LEP students how to solve the kinds of problems that occur on the CTBS. Consistent with this goal, the program is designed to help students to reach the following objectives:

The students can delineate some reasons why people have trouble scoring well on tests.

The students can describe some major features of standardized tests.

The students can mark their answer sheets (or test booklets) quickly and properly.

The students understand the importance of paying attention to test directions.

The students understand the directions on the Reading Vocabulary Test.

The students will check their answers periodically to determine whether they are being marked in the correct spaces.

The students will write down the questions (on scratch paper) that they can't answer quickly, or those which could use further consideration, to assure easy relocation.

The students will go back over the answers to the problems to which they have some uncertainty if time remains.

The students can eliminate answers known to be wrong and guess from the remaining choices if they do not know the answer.

The students can understand the directions on the Reading Comprehension Test.

The students can identify the six types of questions asked on the Reading Comprehension Test and the best technique for answering these questions.

The students can understand the directions on the Spelling test.

The students can implement the strategy for answering Spelling items quickly.

The students can understand the directions for the Language Mechanics test.

The students can implement the strategy for answering Language Mechanics items quickly.

The students can understand the directions on the Language Expression test.

The students can implement the strategies for answering the Language Expression items.

The students can understand the directions for the Mathematics Computation Test.

The students will put scratch paper, if used, under or next to the problem in the booklet and not copy the whole problem onto the scratch paper.

The students will work those problems first which can be done easily and quickly.

The students can understand the directions for the Mathematics Concepts and Applications Test.

The students understand that word problems require careful reading and hard words usually can't be skipped because their meaning is important for understanding the problem.

The students understand directions for Reference Skills Test.

Teaching to the test

Professional ethics demand that we not use for practice any of the problems that are found on the CTBS. That would be teaching the test. One should distinguish, however, between teaching the test and teaching to the test. Teaching to the test involves inferring the objective underlying a given problem, then constructing items for practice that have the same form as those that occur on the test. Use of such items teaches students how to solve problems like those that are found on the test. That is something that is ethical, legal, and desirable. The practice problems in the present program were developed consistent with that principle.

It is suggested that one go even further, presenting the practice materials in exactly the same format as the format of the test. This might include not only the layout of the problems on the page but also the layout and use of separate answer sheets. What one wants to do is to reduce to zero the probability that a student will be intimidated or confused by the stimulus and/or response format of the CTBS. By doing this, one increases the likelihood that the student will be able to tell you what you wanted to know in the first place, that is, whether or not s/he knows how to solve problems of particular types.

Materials

The materials consist of this manual plus a full set of student practice materials. The practice materials are currently available in both English and Spanish.

The manual has five chapters, each chapter providing content for one or more sessions. Each chapter begins with a list of the chapter objectives. The remainder of each chapter is divided into sub-sections, each sub-section corresponding to a single objective. Within each sub-section, there is first a list of the major ideas that are to be developed in relation to the given objective. Following this list is a set of activities designed to help students to accomplish the objective. Each activity is described in sufficient detail to permit the teacher, regardless of his/her training in tests and measurements (or lack thereof), to conduct the activity successfully.

Planning

Q. Who should receive the training outlined in this manual?

A. The training was piloted successfully with fourth, sixth, and eighth grade students. Since the training given is appropriate for use with CTBS Levels 1-4, the effective operating range is fourth grade through twelfth grade.

While the original target population was Spanish-speaking LEP students, the program would be appropriate for use with any English or Spanish speaker who needs assistance in learning how to deal effectively with the CTBS.

Q. When should the training be given?

A. It is best to give the training such that it is completed only shortly before the administration of the CTBS.

Q. How should training be paced?

A. A variety of options are possible running from "massed" practice to "spaced" practice. In one school, eighth grade students completed the entire training program in one day (i.e., massed practice). In another school, fourth grade students completed the training in four sessions spread over four days (i.e., spaced practice). The actual pace you follow will depend on your sense of how much sustained attention your students are capable of giving and how much practice they need. It is recommended that, in most cases, the training be spaced over several sessions on several days. The important thing is that the students get enough practice to actually master the objectives.

Providing for additional practice

You may feel that some students need more practice than is provided by the practice exercises included in this program. In such a case, the given practice exercises provide a model for the construction of additional exercises.

TABLE OF CONTENTS

	Page
SESSION 1	1
Session objectives	1
Objective 1	1
Objective 2	2
Objective 3	4
Objective 4	5
SESSION 2	7
Session objectives	7
Objective 1	7
Objective 2	8
Objective 3	9
Objective 4	10
SESSION 3	13
Session objectives	13
Objective 1	13
Objective 2	14
SESSION 4	17
Session objectives	17
Objective 1	17
Objective 2	18
Objective 3	19
Objective 4	20
Objectives 5 and 6	21
SESSION 5	23
Session objectives	23
Objective 1	23
Objective 2	23
Objective 3	24
Objective 4	24
Objective 5	25
Objective 6	25

SESSION 1

Objectives

1. The students can delineate some reasons why people have trouble scoring well on tests.

2. The students can describe some major features of standardized tests.

3. The students can mark their answer sheets (or test booklets) quickly and properly.

4. The students understand the importance of paying attention to test directions.

Session One

I. Objective: The students can delineate some reasons why people have trouble scoring well on tests.

A. General Ideas

1. Soon we will be taking a test called the Comprehensive Tests of Basic Skills, the CTBS for short.

2. This test is given to students all over the country.

3. I am going to tell you about the kinds of questions you will be asked to answer on the test. This will help you understand exactly what you will have to do before you take the test.

4. We also will spend time learning about and practicing the kinds of skills, techniques, and tricks that will help you improve your scores on the test.

5. Many people do not do well on tests, even when they really know the information, because of a variety of reasons. Some of these reasons are: nervousness, not understanding what the directions mean, working too slowly, marking the answer in the wrong place, etc.

B. Activities

1. Ask students to share with the class how they feel about taking tests and the reasons they feel this way. This activity intends to show students that many, if not most, people have concerns about test-taking.

2. Put a list of reasons on the chalkboard of why many people do not do well on tests (e.g., too nervous, don't understand directions, work too slowly, etc.). Ask students to supplement the list on the board with their own reasons. Discuss each reason and ask students to think about which reasons may apply to them. Tell the students that you will be teaching them strategies which should help them in overcoming many of these problems people have with taking tests.

II. Objective: The students can describe some major features of standardized tests.

A. General Ideas

1. The test that you will be taking, the CTBS, is called a standardized test.

2. A standardized test tests certain skills of yours, like reading, spelling, and math, and your score tells you how well you can perform these skills compared to other people in the country your same age.

3. The CTBS, the test you will be taking, will have questions on it that you are not expected to know the answers to. You should not be bothered by the fact that you do not know the answers to all the questions on the test.

4. To do well on a standardized test, you do not have to answer all of the questions correctly.

5. What is important is that you try to do your best and work as quickly as possible.

6. Your score on the CTBS will not affect your grade in school but it will become a part of your permanent school record.

B. Activities

1. This activity should help students remember some of the features of standardized tests. Put the following true-false questions on the chalkboard.
   (a) Your score on a standardized test tells you how well you perform certain skills compared to other students in the country. (true)
   (b) You are expected to know the answers to all the questions on a standardized test. (false)
   (c) Your score on the CTBS will not affect your grade in school. (true)
   (d) In order to receive a high score on a standardized test, you must answer all the questions correctly. (false)

   Have the students write the answers to these questions on a sheet of paper. Ask for student volunteers to tell whether the statements on the chalkboard are true or false. Discuss each answer with the class.

2. It may be helpful to conduct a discussion with the class differentiating standardized tests from the tests they regularly take in class. The major differences that should be emphasized are (#1):

Standardized tests	Regular class tests
(a) May include material never taught in school.	(a) Usually only includes material taught by the teacher.
(b) Score shows how well one performed on the test compared to others in the country.	(b) Score shows only how well one performed on the test compared to others in the class.
(c) Score does not affect grade in school.	(c) Score does affect grade in school.
(d) You are not expected to know all the answers to the questions on the test.	(d) You are expected to know all the answers to the questions on the test.

   Tell the students to listen and read carefully because you will be giving them a "true/false" test immediately (#2).

III. Objective: The students can mark their answer sheets quickly and properly.

A. General Ideas

Note: Students taking CTBS Level 1 will be marking their answers in the test booklet. Students taking CTBS Level 3 will be using separate answer sheets. The way answer choices should be marked is the same whether they are marked in the test booklet or on a separate answer sheet.

1. Most standardized tests are scored by machine. The machine cannot tell the difference between your answers and stray marks.

2. Mark only one answer for each question.

3. Make your marks dark and clear.

4. Make one heavy dot or line in the center of your answer circle. It will be easier to do this if you use a dull pencil.

5. Do not spend a lot of time blackening in the answer circle because it is important to work fast on a test like this.

6. The whole answer circle does not have to be blackened in for the machine that scores these tests to pick it up.

7. Do not make marks anywhere else but in the answer circle. Accidental marks may cause an answer to be counted wrong.

8. If you make a mistake or want to change an answer, erase your first answer completely. If you leave a little bit of pencil mark when you erase the first answer you chose, the machine that marks your paper will not be able to tell which answer you want. It will mark the question wrong.

B. Activities

1. Draw a sample answer set on the chalkboard like the following:
1. ① ② ③ ④
Demonstrate to students how to fill in the answer circle and how a changed answer should be completely erased.

2. Pass out the Practice Answer Sheets (#3) to the class so that the students can practice marking the answer choices quickly and properly. On the Answer Sheets, there are 10 items and 4 answer choices for each item. Call out the number of each item and the number of the answer choice students should mark for each item. For example, question 1, circle 3; question 2, circle 5; question 3, circle 4; question 4, circle 8; question 5, circle 2; question 6, circle 6; question 7, circle D; question 8, circle E; question 9, circle A; question 10, circle G. Check to see that students are marking their answers with quick pencil strokes and not spending time filling in the circle completely.

3. This activity is designed so that students can identify answer circles that are improperly marked. Put the following on the chalkboard.

Ask students to write down what the problem is with the way the above answer choices have been marked. Have student volunteers describe to the class what they think the problem is with the way each of the answer choices has been marked.

IV. Objective: The students understand the importance of paying attention to directions.

A. General Ideas

1. It is extremely important to read each word in the directions very carefully and listen carefully to the directions given orally by the teacher.

2. The directions tell you about the kinds of questions you will be answering and the basis on which you should select your answer.

3. If you do not follow the directions you will most likely lose many points.

4. Do not assume you know what to do without reading the directions carefully and listening to the person giving the test.

5. If you do not understand the directions, ask the person giving the test to explain anything that is unclear.

6. Keep the directions in mind while taking the test.

B. Activity

This activity focuses on the importance of following directions. Distribute to the class the practice exercise labelled Directions (#4). Tell the students that this exercise is a special set of tasks designed to see how good they are at following directions.

SESSION 2

Objectives

1. The students understand the directions on the Reading Vocabulary Test.

2. The students will check their answers periodically to determine whether they are being marked in the correct spaces.

3. The students will write down the questions (on scratch paper) that they can't answer quickly, or those which could use further consideration, to assure easy relocation.

   The students will go back over the answers to the problems to which they have some uncertainty if time remains.

4. The students can eliminate answers known to be wrong and guess from the remaining choices if they do not know the answer.

Session Two

I. Objective: The students understand the directions on the Reading Vocabulary Test.

   A. General Ideas

      1. There are eight different sections of the CTBS, each of them is called a test.

      2. We will go over the directions for each test so that you will understand what you are supposed to do.

      3. Often people taking tests will miss questions, not because they do not know the answers, but because they do not understand what they are supposed to do.

      4. Be on the lookout for the important words in the directions which tell you the basis on which you should answer the questions (e.g., Choose the word that means the same as ...; Choose the best answer ...). These are the words you have to pay attention to and ask the teacher for clarification if you do not understand what they mean.

5. The first section of the CTBS is called the Reading Vocabulary Test and the directions say to "Choose the word that means the same, or about the same, as the underlined word." We will discuss very specifically what these directions mean.

B. Activities

1. This activity helps students understand what directions intend when they say "Choose a word that has the same, or about the same, meaning as another word. Put the words "big" and "large" on the chalkboard. Tell the students that these are words that have about the same meaning. This means that these words can be used interchangeably in sentences. Have students generate sentences using the words "big" and "large." Put other words on the chalkboard (e.g., fast, cold) and ask students to use these words in sentences and to think of other words that can be substituted for them but where the original meaning is retained.

2. This activity helps students understand what it means to select the word that has the closest meaning to the underlined word from a group of four words. Put the following example on the chalkboard:

   tiny babies
   (a) small
   (b) weak
   (c) helpless
   (d) crying

   Tell the class that although all the answer choices could be used before the word "babies" only one of the choices has almost the same meaning as the word "tiny." Ask the students to select the word that means about the same thing as the word "tiny." Go over the correct answer with the class explaining, if necessary, why the other choices are incorrect. Another example that can be used for this exercise is:

   fresh strawberries
   (a) frozen
   (b) red
   (c) ripe
   (d) just picked

II. Objective: The students will check their answers periodically to determine whether they are being marked in the correct spaces. (Note: This objective is only relevant when separate answer sheets are used.)

A. General Ideas

1. When there are questions that you cannot answer quickly you should write down the number of the question on scratch paper so that you can return to it if you have time.

2. If you do skip a question and plan to return to it later, you will have to remember to skip a space on your answer sheet.

3. Students sometimes forget to skip a space on their answer sheets and consequently mark their answers in the wrong spaces. When this happens, their answers will be counted as incorrect.

4. There are strategies to help make sure you have marked your answers in the correct spaces.
   (a) Check frequently that the number of the question you are marking on your answer sheet corresponds to the question you are answering in the test booklet.
   (b) Some people find it helpful to vocalize the question number and the answer number as they look from the test booklet to the answer sheet.

III. Objectives: Students will write down on scratch paper the number of those items they cannot answer quickly, or those which could use further consideration, to assure easy relocation.

The students will go back over the answers to the problems to which they have some uncertainty if time remains.

A. General Ideas

1. Standardized tests have hard and easy questions.

2. Since standardized tests are timed, you have to work as quickly as possible.

3. If you spend too much time on a hard question, you may not have time to finish the test. And, some of the questions you did not have time to answer may have been very easy.

4. It is better to answer all the questions first that you can answer quickly. These are the questions you know the answer to or can make a quick guess. We will learn how to make good guesses later.

5. When you come to a question that you feel you cannot answer quickly, write the number of that question down on your scratch paper.

6. When you come to the end of a section (or the end of the test) and still have some time remaining, you can look at the numbers on your scratch paper to find out which questions you need to go back and work on.

7. You can also mark down questions that you want to spend more time thinking about.

8. I will be reminding you throughout the lesson to check your answers, particularly those you are uncertain of, if you finish early.

IV. Objective: The students can eliminate answers known to be wrong and guess from the remaining choices if they do not know the answer.

The students know that they should guess when they do not know the correct answer.

A. General Ideas

1. On those questions where you don't know the correct answer, you should guess at an answer rather than leave it blank.

2. Sometimes test directions will tell you not to guess because points will be subtracted for wrong answers. We will be discussing how to make good guesses so that it is always better to guess than to leave the question blank.

3. When you are not sure about the right answer you will have to guess in order to answer the question.

4. The first thing to do when guessing is to decide which answers could not be right.

5. From the answers that are left, guess which one you think might be right and mark that answer choice on your answer sheet.

6. Mark down the number of the question you answered by guessing so that if you have time at the end of the test, you can spend time trying to figure out the correct answer.

7. Do <u>not</u> leave any questions unanswered.

B. Activities

1. Lead the students to consider the probabilities associated with chance guessing. They should come to understand that the chances for guessing correctly improve dramatically as one is able to reduce the number of options among which one has to choose. You might approach this by reminding them of the TV game show format where the contestant must guess among several doors one of which hides a prize. If there are four doors, then the probability of winning the prize is 1:4 or 25%. If we have inside information that allows us to eliminate one door, the probability is improved to 1:3 or 33⅓%. If we have additional information that allows us to eliminate another door, the probability is 1:2 or 50%.

2. The following activity is designed to help students understand how to eliminate wrong answers first. Tell students the answer choices are one, two, three, and four and you know that two and three are not the correct answers. Ask students which answers they should guess between (i.e. one and four). Continue giving students problems like these. Tell them that when they are able to eliminate at least one answer choice, it makes it more likely they will choose the correct answer.

3. Put the following examples on the chalkboard and ask the class to try to eliminate incorrect answer choices and guess at the correct answer. Have student volunteers tell how they can eliminate certain answer choices even if they are not sure of the correct answer.

   e.g., Give means the same as
   (a) invest
   (b) go
   (c) contribute
   (d) steal

   e.g., Which word is not spelled correctly
   (a) camel
   (b) pig
   (c) girafe
   (d) donkey

4. Ask students to articulate the rules for answering questions on a test.
   (i.e., 1. Answer all questions first that can be answered quickly.
   2. Check periodically to make sure you are marking your answers in the right place.

3. Write the numbers of the questions on scratch paper that require more time to answer and return to these later.
4. If you do not know an answer, guess.)

By articulating these rules, it will help students remember them.

5. Have the students do Practice Exercise #5a or #5b. Tell them that this is like the Reading Vocabulary Test but that it is designed to give them practice in doing some of the things you have been talking about. Go over the instructions on the facing page. Then monitor their work.
Make sure that:
   a. they are marking the easy items first
   b. they are marking their answers in the right places
   c. they are writing the numbers of omitted items on scratch paper.

   Once they have finished, correct the items but not by merely giving the correct answer. Do so rather by eliminating incorrect answers, thus illustrating the guessing procedures that you have been teaching.

6. Distribute the Practice Exercises -- Reading Vocabulary Test (#6a and #6b) to the students. The purpose of this activity is twofold: 1) to make sure the students understand the test directions, and 2) to make sure the students periodically check their answers to determine if they are marked in the correct spaces.

   Ask students to read the directions to themselves and look over the sample item(s). Then ask a student volunteer to tell the class what the directions mean. Emphasize to the class that the most important part of the directions is the sentence that says "Choose the word that means the same, or about the same, as the word with the line under it."

7. Ask students to complete the practice exercises. Remind students of the following information: 1) They must work as quickly as possible; 2) If there are items that they cannot answer quickly, they should write their numbers down and do them last, after they have completed the easier items; 3) They should check regularly to see whether they are marking their answers in the correct spaces; 4) If there is time remaining, they should go back and check their answers; and 5) If they do not know an answer, they should guess.

   As the first students complete the exercises, let the rest of the class know they have only a minute left. Go over the correct answers with the students.

SESSION 3

Objectives

1. The students can understand the directions on the Reading Comprehension Test.

2. The students can identify the six types of questions asked on the Reading Comprehension Test and the best technique for answering these questions.

Session Three

I. Objective: The students can understand the directions on the Reading Comprehension test.

A. General Ideas

1. The second section of the CTBS is called Reading Comprehension.

2. The Reading Comprehension section is designed to show how well you understand what you read.

3. You must read a short selection (e.g., a story, poem, etc.) and then answer questions about it.

4. On this test the directions say to choose the "best" answer. The "best" answer is the one that most accurately describes what occurred in the selection.

B. Activities

1. Distribute an example of the type of item encountered in the "Reading Comprehension" test (#7). Point to the reading selection and to the item based on it.

2. Tell students the following information. On the Reading Comprehension Test, they are to choose the best answer for each question. The best answer is the one that most accurately answers the question. Sometimes more than one answer seems like it could be correct, but in the Reading Comprehension test, the best answer is the one that most accurately represents what is contained in the reading selection.

Using the example you have distributed, demonstrate

how several of the answer choices could be a correct (i.e., plausible) answer to the question, but that only one accurately represents what was described in the selection.

II. Objective: The students can identify the six types of questions asked on the Reading Comprehension Test and the best technique for answering these questions.

A. General Ideas

1. There are basically six different kinds of questions that are asked on the Reading Comprehension Test.

2. These six different kinds of questions ask for the following:
   (a) details
   e.g., What happened? Why did _____ happen? When did _____ happen? Who did _____?
   (b) order in which events occur
   e.g., What happened first?
   (c) meaning of words from context
   e.g., What does the word _____ in line __ mean?
   (d) using descriptive words properly
   e.g., Which word best describes _____?
   (e) making inferences
   e.g., What does _____ show? What is the most likely cause of _____? Why did _____ probably happen?
   (f) main idea or purpose
   e.g., What is this story mainly about? What is the best title?

3. Answering the different kinds of questions on the Reading Comprehension Test requires the use of different strategies.

4. It is generally a good idea to first read the selection completely through as quickly as you can.

5. Then try to answer as many questions as you can.

6. If you do not know the answer to a question that asks about <u>details</u> or the <u>order in which events occur</u> go back only to that part of the story that will help you answer the question.
   -- Quickly glance through the story until you come to the part that answers the question.
   -- Do <u>not</u> read the whole story again.
   -- Do <u>not</u> spend time trying to figure out words you do not know unless they are important for

answering the question.
-- Detail and order questions can be answered even if you do not understand many aspects of the selection.

7. In order to answer meaning of words from context questions, you should locate the line(s) in the selection where the word is used.
   -- Do not read the whole story again.
   -- Figure out which answer choice is closest in meaning to the way the word is used in the selection.
   -- All of the answer choices may be correct uses of the word, but only one answer choice depicts the way the word is used in the selection.
   -- Meaning of words from context questions may be answered knowledgeably even if there are other aspects of the selection that you do not understand.

8. Answering descriptive words questions requires understanding how people in the selection feel, how things look, etc. and choosing the answer choice that most accurately depicts this understanding.
   -- Descriptive words questions require a careful reading of that portion of the selection where the answer is described.
   -- There is no need to carefully reread the entire selection.
   -- The correct answer choice uses words that are different from those used in the selection, but mean about the same.

9. Making inferences questions require you to draw conclusions based on what you read.
   -- The answer will not directly appear in the selection.
   -- Generally you will have to have a fairly good understanding of the selection.
   -- The correct answer choice is the one which is the most likely conclusion based on the relevant information in the selection.

10. Main idea or purpose questions require an understanding of the point of the selection.
    -- The answer will not directly appear in the selection.
    -- You will need to have a good understanding of the entire selection.
    -- However, you do not need to know the meaning of every word.

B. Activities

1. Distribute Practice Exercises -- Reading Comprehension (#8a and #8b). Ask students to read the first selection. Go over each question with them helping the students determine what kind of information the question asks for and the type of strategy which should be used. Have students determine which answer choices are correct and go over the answers with them.

2. Students then should read the second selection and mark the answer choices which they believe are correct. Go over correct answers with the class.

SESSION 4

Objectives

1. The students can understand the directions on the Spelling test.

2. The students can implement the strategy for answering Spelling items quickly.

3. The students can understand the directions for the Language Mechanics test.

4. The students can implement the strategy for answering Language Mechanics items quickly.

5. The students can understand the directions on the Language Expression test.

6. The students can implement the strategies for answering the Language Expression items.

Session Four

I. Objective: The students can understand the directions on the Spelling test.

A. General Ideas

1. This is CTBS Test 3 -- Spelling.

2. Each item is a sentence with a word or words underlined.

3. You must decide which word, if any, is misspelled.

B. Activity

Distribute the Practice Exercise -- Spelling to the class. Go over the directions to this test with the class using the sample item(s) as an example of what the directions mean.

II. Objective: The students can implement the strategy for answering spelling items quickly.

A. General Ideas

1. Look at the underlined word(s) first.

2. If you know the word(s) is spelled incorrectly, blacken the correct space. There is no need to read the sentence.

3. If the word(s) looks like it has been spelled correctly, read the sentence to see if the correct word has been used in the sentence.

4. The spelling errors in the sentences are of two kinds: (a) Sometimes words that sound the same, but are spelled differently and have different meanings are used instead of the correct word. (b) Sometimes words that are phonetically similar (i.e., sound the same, but are spelled differently) to the correct word are used. These phonetically similar words actually are not words at all.

B. Activities

1. Give students practice in selecting the correctly spelled word with pairs of phonetic equivalents. The following are some phonetic equivalents to use for practice: after - aftre, cri - cry, candy - candey, snow - snoe, coat - cote, fil - fill. Put these phonetic equivalents on the chalkboard and ask students to write down the one in each pair which is spelled correctly. Ask student volunteers to offer the correct answers.

2. Students who are clear about the meanings of words that sound the same but are spelled differently (i.e., homonyms) should have an easier time with the spelling tests. Give students practice in determining the meanings of various homonyms. The following are some examples of homonyms to use for practice: blue - blew, there - their, too - two, no - know, by - buy. Put these homonyms on the chalkboard. Ask students to write a sentence using each word and write the meaning of each word. Student volunteers can offer the correct answers.

3. Have students finish the Spelling Practice Exercises. Remind them to work quickly. They should first look

at the underlined word(s) without reading the sentence. If they know an underlined word is spelled incorrectly, there is no need to read the entire sentence. They should simply mark the correct answer choice. If the underlined word(s) seems to be spelled correctly, the entire sentence should be read to determine if the homonym with the proper meaning is used in the sentence.

Tell students to remember to guess if they are not sure of the correct answer.

III. Objective: The students can understand the directions on the Language Mechanics test.

A. General Ideas: Punctuation

1. We will be going over the directions for the CTBS Language Mechanics Test.

2. There are two sections to this test, punctuation and capitalization.

3. In the punctuation section, you are to determine whether the sentence is punctuated correctly.

4. If the sentence is punctuated correctly you are to choose the answer choice "none" (for Level I) or "best as it is" (for Level III).

5. If the sentence is punctuated incorrectly, choose the answer with the punctuation that would correctly punctuate the sentence.

B. Activity

Distribute the Practice Exercises -- Language Mechanics to the students. After reading or listening to the directions on the punctuation section, have student volunteers explain how the sample item should be answered according to the directions. After the sample item has been answered correctly, put the sample item on the chalkboard punctuating it correctly. (i.e., He plays basketball, football, and baseball.) Ask students which answer choice they should mark if the item were written in this way. Go over the correct answer with the class (i.e., for Level I "none"; for Level III "best as it is").

IV. Objective: The student can implement the strategy for answering Language Mechanics items quickly.

A. General Ideas: Punctuation

1. Do not spend time trying to figure out the meaning of every word in the sentence.

2. Look for cues that indicate punctuation errors (i.e., dates without commas, abbreviations without periods).

3. Even if there are words you do not know, it is likely you will be able to figure out whether the sentence is punctuated properly.

4. This same strategy will be used on the Capitalization section, only capitalization errors rather than punctuation errors are the focus.

B. Activity

Students should answer the punctuation items on the Practice Exercises -- Language Mechanics implementing the strategy for answering the items quickly. Go over the correct answers.

C. General Ideas: Capitalization

1. In the Capitalization section you are to decide where a capital letter is needed.

2. Each of the sentences in this section is divided into three parts.

3. If a word in any part needs a capital letter and isn't capitalized, for Level I, fill in the answer space under the part where the capital is needed; for Level III, fill in the space on your answer sheet that has the same numeral as the part where the capital is needed.

4. Remind students to use the same strategy for answering these items as they did on the punctuation section.

D. Activities and Exercises

1. Distribute the capitalization items on the Practice Exercises -- Language Mechanics. Go over the

directions for this section using the sample item as an example. Have students complete exercises reminding them to look for capitalization errors before reading the sentence for understanding.

V & VI. Objectives: The students can understand the directions for the Language Expression Test.

The students can implement the strategies for answering the Language Expression items.

A. General Ideas

1. The Language Expression Test shows how well you use the English language to express yourself.

2. There are six sections of the Language Expression Test and the directions are different for each section. This means you have to pay careful attention.

3. There are several strategies for answering the various Language Expression items.
   (a) Some items call for you to choose the most grammatically correct use of the English language. With these items, it is a good idea to vocalize to yourself the item and answer choices so you can try to hear the answer that sounds correct. If you remember the relevant grammar rule, that may help you select the correct answer.
   (b) Some items ask you to determine the correct meaning of words in context or select the most appropriate word to fit a sentence. These items require you to understand the context in which a word is used. Read the passage and see what the meaning of a specified word is in that passage. If you are to select the best word for a sentence, read the sentence carefully to determine which word fits the meaning of the sentence most precisely.
   (c) Some items require you to determine the proper order of sentences as they should appear in a paragraph. With Level I, students are to determine which sentence should be first. With Level III, students are to decide the proper order for all the sentences. With these items, the meaning of each item is important for determining the proper sequential order. Some words (e.g., Then, But, And) which start sentences are cues that these sentences are not the first ones in the paragraphs.

B. Activity

Distribute the practice exercises to the class. Go over the directions and sample items for each section. Once the students can articulate what the directions mean for a section, tell them the strategy for answering those items. Let students practice answering the questions using the strategy. Remind students to work quickly, and guess by eliminating incorrect answer choices when they are not sure of the correct answer.

SESSION 5

Objectives

1. The students can understand the directions for the Mathematics Computation Test.

2. The students will put scratch paper, if used, under or next to the problem in the booklet and not copy the whole problem onto the scratch paper.

3. The students will work those problems first which can be done easily and quickly.

4. The students can understand the directions for the Mathematics Concepts and Applications Test.

5. The students understand that word problems require careful reading and hard words usually can't be skipped because their meaning is important for understanding the problem.

6. The students understand directions for Reference Skills Test.

Session Five

I. Objective: The students can understand the directions on the Mathematics Computation Test.

   A. General Ideas

      1. The directions on the Mathematics Computation Test are straightforward.

      2. There are four sections to this test: addition, subtraction, multiplication, and division.

      3. It is important to keep in mind the operation (i.e., add, subtract, etc.) required for each section. Sometimes people forget they are working on a subtraction section, for example, and start adding. This will lead to a wrong answer. Check yourself to make sure you have performed the proper operation.

II. Objective: The students will put scratch paper, if used, under or next to the problem in the booklet and not copy the whole problem onto the scratch paper.

A. General Ideas

1. There are many problems in the Mathematics Computation section that you will not be able to figure out in your head. For this reason, you are allowed to use scratch paper.

2. When you use scratch paper, do not waste your time copying the entire problem down. The way to avoid having to copy the entire problem down is to put the scratch paper next to or under the problem you are working on in the test booklet.

III. Objective: The students will work those problems first which can be done easily and quickly.

A. General Ideas

1. The problems in each section (i.e., addition, subtraction, etc.) of the Mathematics Computation Test seem to increase in difficulty as the section progresses.

2. Therefore, it may be better to do those problems for each section first which can be done quickly and easily and then go back and finish the rest.

B. Activities

1. Put the following strategy reminder on the chalkboard.
   Mathematics Computation
   (a) Check operation.
   (b) Don't copy entire problem on scratch paper.
   (c) Work easy problems first.

   Ask students to articulate what each reminder means.

2. Distribute the Practice Exercises to the class. Have students complete the exercises keeping the strategy reminders in mind.

IV. Objective: The students can understand the directions on the Mathematics Concepts and Application Test.

A. General Ideas

1. The Mathematics Concepts and Application Test shows how well you can use math words and concepts and do word problems.

2. Each item asks a question and you are to select the answer choice that correctly answers the question.

V. Objective: The students can understand that the problems in this section are word problems and must be read carefully.

A. General Ideas

1. Word problems must be read carefully.

2. In these problems, most of the hard words cannot be skipped because they are important in determining which answer choice is correct.

3. The mathematics in this section is usually not too difficult, but the mathematics vocabulary may be problematic.

4. If you do not understand all the words in the question, try to determine what the important words must mean. Then eliminate the answer choices that seem like they provide the wrong answer, and choose the answer choice that seems like it is correct.

B. Activities

1. Distribute the Practice Exercises to the class. Explain to the class that these exercises will give them an idea of the type of questions that will be on the Mathematics Concepts and Application Test. But, also explain that there will be many more problems than those which appear on the Practice Exercises.

2. Go over the directions with the class using the sample item as an example. Have students complete the exercises. Have student volunteers give the correct answers to the problems and share with the class how they decided on their answers.

VI. Objective: The students can understand the directions for the Reference Skills Test.

A. General Ideas

1. The items in the Reference Skills section are about finding information in the library and in books.

2. You are to choose the correct answer for each item.

REFERENCES

Callenbach, C. The effects of instruction and practice in content-independent test-taking techniques upon the standardized reading test scores of selected second grade students. Journal of Educational Measurement, 1973, 10, 25-30.

Cummins, J. The entry and exit fallacy in bilingual education. Paper presented at the ninth annual meeting of the National Association for Bilingual Education, San Diego, April 1980.

Cummins, J. Language proficiency and academic achievement. In J. W. Oller, Jr. (Ed.), Current Issues in Language Testing Research, Rowley, Mass.: Newbury House (in press), 1981.

Diamond, J. J. and Evans, W. J. An investigation of the cognitive correlates of test-wiseness. Journal of Educational Measurement, 1972, 9, 145-150.

Ebel, R. L. and Damrin, D. E. Tests and examinations. In C. W. Harris (Ed.), Encyclopedia of Educational Research, 3rd Edition, N. Y.: The Mac Millan Company, 1960.

Ford, V. A. Everything you wanted to know about test-wiseness. Princeton, N. J.: Educational Testing Service, 1973 (ERIC Document Reproduction Service No. ED 093912).

Helwig, J. T. and Council, K. A. (Eds.), The SAS User's Guide, 1978 Edition, SAS Institute, Cary, North Carolina, 1979.

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole Publishing Co., 1968.

Langer, G., Wark, D., and Johnson, S. Test-wiseness in objective tests. In P. L. Nacke (Ed.), Diversity in mature reading: Theory and research. Vol. 1, 22nd Yearbook of the National Reading Conference. National Reading Conference, Milwaukee, Wisc., 1973.

Millman, J. Test-wiseness in taking objective achievement and aptitude examinations. Final Report, 1966, College Entrance Examination Board.

Millman, J., Bishop, C. H. and Ebel, R. An analysis of test-wiseness, Educational and Psychological Measurement, 1965, 25, 707-726.

Moore, J. C., Schutz, R. E. and Baker, R. L. The application of a self-instructional technique to develop a test-taking strategy. American Educational Research Journal, 1966, 3, 13-17.

Nilsson, I. and Wedman, I. On test-wiseness and some related constructs, Educational Reports, UMEA, No. 7, 1974.

Oakland, T. The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. Journal of School Psychology, 1972, 10, 355-360.

Sarnacki, R. W. An examination of test-wiseness in the cognitive test domain. Review of Educational Research, 1979, 49, 252-279.

Slakter, M. J., Koehler, R. A. and Hampton, S. H. Learning test-wiseness by programmed texts. Journal of Educational Measurement, 1970, 7, 247-254.

Thorndike, E. L. Reliability. In E. F. Lindquist (Ed.), Educational Measurement. Washington, D. C. American Council on Education, 1951.

Wahlstrom, M. and Boersma, F. J. The influence of test-wiseness upon achievement. Educational and Psychological Measurement, 1968, 28, 413-420.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Woodley, K. K. Test-wiseness: A cognitive function? Paper presented at the annual meeting of the National Council on Measurement in Education, Washington, D. C., April 1975.